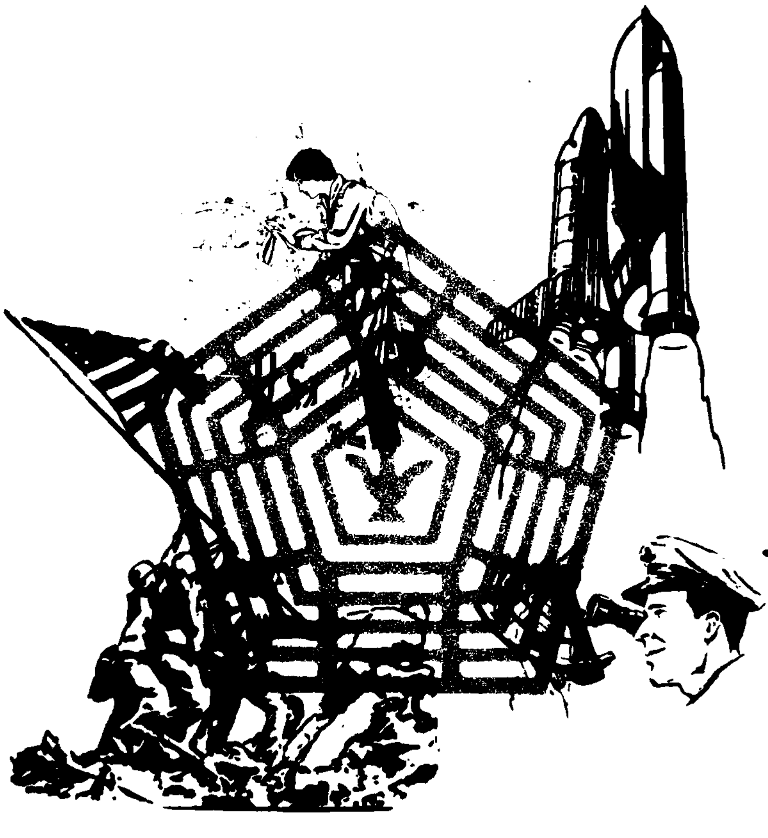
USAFA-TR-84-2

(12)

PROCEEDINGS
PSYCHOLOGY IN THE DEPARTMENT OF DEFENSE

AD-A141 043



NINTH SYMPOSIUM

18 - 20 APRIL 1984

SPONSOR:
USAF ACADEMY DEPARTMENT OF BEHAVIORAL SCIENCES AND LEADERSHIP

84 05 01

PROCEEDINGS OF THE NINTH SYMPOSIUM

PSYCHOLOGY IN THE DEPARTMENT OF DEFENSE

18 - 20 April 1984

Department of Behavioral Sciences and Leadership

United States Air Force Academy

Colorado Springs, Colorado

Symposium Chairman

Major Frank R. Wood

EDITORS

Captain George E. Lee
Major Thomas E. Ulrich

SYMPOSIUM COMMITTEES

Professor and Acting Head, Department of
Behavioral Sciences and Leadership

Colonel William E. Rosenbach

EXECUTIVE COMMITTEE
Major Frank R. Wood, Chair
Colonel William E. Rosenbach
Lieutenant Colonel John E. Anderson
Major William H. Clover
Major Robert A. Gregory
Major Thomas M. McCloy
Major William L. Derrick

PROGRAM COMMITTEE
Major Thomas E. Ulrich, Chair
Captain William H. Cummings III
Captain Gregory D. Smith
Captain Kevin W. O'Callaghan
Captain June I. Babson
Captain George E. Lee

INVITATIONS AND ATTENDANCE COMMITTEE
Captain William P. Marshak, Chair
Captain Paul M. Grunzke
Captain James A. Boyless

SOCIAL, BILLETING & TRANSPORTATION COMMITTEE
Major Dean H. Wilson, Chair
Captain Milagros Caraballo
Captain George W. Yeakey
Captain Michael A. Rosebush

FACILITIES AND VISUAL AIDS COMMITTEE
Captain Karen O. Dunivin, Chair
Staff Sergeant Robert L. Stewart
Sergeant Larry J. Dannelly

PROTOCOL COMMITTEE
Captain Ruth E. Freese, Chair
Captain Michael N. Stollings
Captain Jose Bolton

PUBLICITY COMMITTEE
Captain Philip A. Irish III, Chair
Captain Paul J. Johnston

This research report is presented as a competent treatment of the subject, worthy of publication. The United States Air Force Academy vouches for the quality of the research, without necessarily endorsing the opinions and conclusions of the authors.

This report has been cleared for open publication and/or public release by the appropriate Office of Information in accordance with AFR 190-17 and DODD 5230.0. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service.

This research report has been reviewed and is approved for publication.

Thomas E. McCann

THOMAS E. McCANN, Lt Col, USAF
Director of Research, Studies
and Analysis

Additional copies of this document are available through the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22151.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: USAFA TR 84-2
2. GOVT ACCESSION NO.: ADA
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE *(and Subtitle)*: Proceedings, Psychology in the Department of Defense, Ninth Annual Symposium
5. TYPE OF REPORT & PERIOD COVERED: Proceedings
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s):
   Captain George E. Lee (Ed.)
   Major Thomas E. Ulrich (Ed.)
8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS:
   Department of Behavioral Sciences and Leadership
   U.S. Air Force Academy
   Colorado Springs CO 80840
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS:
   HQ USAFA/DFBL
   USAF Academy
   Colorado Springs CO 80840
12. REPORT DATE: April 1984
13. NUMBER OF PAGES:
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*:
15. SECURITY CLASS. *(of this report)*: Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for Public Release and Distribution Unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Psychology	Radiation Effects	Assessment	Workload
Symposium	Counseling	Stress	Family Issues
Training	Cognitive Functioning	Job Performance	Biofeedback
Leadership	Operator Performance	Turnover	Retention

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

These printed proceedings include papers and presentation that deal with a wide range of research in psychology with emphasis on military issues.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE iv

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data F

FOREWORD

The opinions expressed in this volume are those of the individual authors and do not necessarily reflect official policy either of the Army, Navy, Air Force , Coast Guard or of any government organization in which the author(s) may be serving.

Trade names of materials or products of commercial or non-government organizations are cited only where essential to precision in describing research procedures or evaluation of results. Their use does not constitute official indorsement or approval of the use of such commercial hardware or software.

These proceedings are approved for public release, distribution unlimited, in accordance with paragraph IVA, AFR 80-45.

In publishing the Proceedings of the Ninth Symposium, the editor sought to facilitate an effective and timely dissemination of the technical information presented. The papers contained in this document were printed directly from unedited reproducible copy submitted by the authors who are solely responsible for their contents.

Acknowledgements

Cover: Graphics Division of Audiovisual Services
Dean of Faculty, USAF Academy

Editorial Assistance:

Helen Wilson
Sharon Weatherson
Janet Gomes

OVERVIEW OF SYMPOSIUM SCHEDULE

Wednesday, 18 April 1984

Morning Activities	Event
	REGISTRATION
	KEYNOTE ADDRESS
	Dr. J. Richard Hackman, Yale University Topic: "Developing usable knowledge about team effectiveness: Requirements for theory and method"
	PANEL DISCUSSION AND PAPER SESSIONS
	Panel: Testing the Human Dimensions of Command, Control, Communications and Intelligence (C^3I) Systems
	Paper Session: Social Control
	Paper Session: Counseling
	Paper Session: Investigating Theories of Leadership
	Paper Session: Effects of Irradiation on Animal Performance
	INVITED LUNCHEON ADDRESS
	Dr. David P. Campbell (Center for Creative Leadership) Topic: "A Personality Profile for General Officers"

Afternoon Activities	Event
	PANEL DISCUSSION AND PAPER SESSIONS
	Panel: Radiation-Induced Combat Performance Degradation: The Defense Nuclear Agency Intermediate Dose Program
	Panel: Physical Performance Tests as Predictors of Task Performance
	Panel: Simulator-Induced Sickness: Reaction to a Transformed Perceptual World
	Paper Session: Job Perceptions, Attitudes, and Performance

Panel: The Effects of Prescribed Medications on Cognitive Functioning

Panel: The Well-Being of Army Soldiers and Their Families: Methodological, Substantive, and Technical Considerations

Paper Session: The Study of Operator Performance

Paper Session: Selection and Placement

Paper Session: Mental Health and Morale

Panel: The Effects of Prescribed Medications on Cognitive Functioning (Part II - Continuation of Workshop from Previous Session)

Thursday, 19 April 1984

Morning Activities	Event

INVITED ADDRESS

Dr. Harry L. Synder (Virginia Polytechnic Institute)

Topic: "Visual Display Technology - A Developing Challenge for the Behavioral Scientist"

PANEL DISCUSSIONS AND PAPER SESSIONS

Panel: A Systems Approach to the Study of Leadership

Panel: Animal Models for the Assessment of High Risk Environments

Paper Session: Causes, Consequence, and Control of Stress

Paper Session: Defining Leadership

Paper Session: The Management of Information in Organizations

SYMPOSIUM LUNCHEON ADDRESS

Colonel John J. Clune (Director of Athletics, U.S. Air Force Academy)

Topic: "Psychology and Big Time Intercollegiate Athletics"

(19 Apr 84 Continued)

Afternoon Activities	Event
	PANEL DISCUSSIONS AND PAPER SESSIONS
	Panel: Coping with Military Personnel Turnover: Some Applied Research
	Panel: Emerging Issues in Aerospace Team Training and Selection
	Panel: Individualized Instruction
	Paper Session: Issues in the Study of Workload
	Paper Session: Family Issues
	Panel: Biofeedback and the Behavioral Sciences
	Panel: Perceptual Factors In Low Altitude Flight
	Panel: Family Support Centers and Their Relation to Quality Force, Mission Support, and Readiness
	Paper Session: Recruiting and Retention of Military Personnel
	Panel: Psychophysiological Tools in Engineering Psychology

Friday, 20 April 1984

Morning Activities	Event
	PANEL DISCUSSIONS AND PAPER SESSIONS
	Panel: Air Force Basic Research on Ability Measurement
	Panel: Man in Space: Space Technology and the Behavioral Aspects
	Paper Session: Tests and Measurement Issues
	Paper Session: Issues in Military Leadership
	Paper Session: The Study of Working Groups and Crews

(20 Apr 84 Continued)

Panel: Defining the Critical Parameters of Computer Literacy

Paper Session: Research on Vision and Perception

Panel: Cognitive Strategies Training

Panel: Technology Transfer: National Security Policy and the Scientific and Academic Community

Paper Session: The Design of Research and Training Programs

PROCEEDINGS OF THE NINTH SYMPOSIUM ON PSYCHOLOGY
IN THE DEPARTMENT OF DEFENSE

WEDNESDAY, 18 APRIL 1984

THURSDAY, 19 APRIL 1984

FRIDAY, 20 APRIL 1984

KEYNOTE ADDRESS

DEVELOPING USABLE KNOWLEDGE ABOUT TEAM EFFECTIVENESS; REQUIREMENTS FOR THEORY AND METHOD

Dr. J. Richard Hackman
(Yale University)

J. Richard Hackman is Professor of Organizational Behavior and of Psychology at Yale University. He received his undergraduate degree in mathematics from MacMurray College in 1962, and his doctorate in social psychology from the University of Illinois in 1966. He has been at Yale since then.

Professor Hackman conducts research on a variety of topics in social and organizational psychology, including the design of work, the task effectiveness of work groups, and social influences on individual behavior. He is on the editorial board of several professional journals, and has consulted with a number of organizations on quality of worklife issues.

He is the author or editor of five books (the most recent being Work Redesign, co-authored with Greg R. Oldham and published in 1980 by Addison-Wesley) and over 50 chapters and articles. Professor Hackman was winner of the Sixth Annual AIR Creative Talent Award in the field of "Measurement and Evaluation: Individual and Group Behavior," and co-winner of the 1972 Cattell Award of the American Psychological Association. He is a Fellow of that association in the Division of Industrial and Organizational Psychology, and in the Division of Personality and Social Psychology.

Dr. Hackman's address at the Ninth Biennial Psychology in the DOD Symposioum on 18 April 1984 is titled "Developing Usable Knowledge About Team Effectiveness: Requirements for Theory and Method." A summary of Professor Hackman's address follows.

DEVELOPING USABLE KNOWLEDGE ABOUT TEAM EFFECTIVENESS:
REQUIREMENTS FOR THEORY AND METHOD

Dr. J. Richard Hackman

Yale University

This paper explores the present state of knowledge about work team effectiveness, and proposes some possible new directions for theory and method in small group research. It is suggested that descriptive models of group behavior that emphasize the role of group interaction process in mediating between "input" and "output" states may have misguided research and theory on the topic. As an alternative, a case is made for attending more carefully to certain aspects of group structure (particularly the authority structure, group composition, and group task design) in group performance research, and for developing theory that illuminates the relationship between a team and the organizational context in which it operates. Regarding methodology, questions are raised about the usefulness of both experimental and survey studies of group effectiveness, in some non-traditional methods for learning about the determinants of group performance are suggested.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 248	Subordinate Perceptions of Contingent Leaders: Do Followers Accept our Theories.
AD-P003 249	To the Wilderness and Beyond: The application of a Model for Transformal Change.
AD-P003 250	Comparative Effects of Bremsstrahlung, Gamma, and Electron Radiation on Rat Motor Performance.
AD-P003 251	The Interactive Effects of Morphine and Ionizing Radiation on the Latency of Tail-Withdrawal from Warm Water in the Rat.
AD-P003 252	Quarternary Naltrexone Reverses Radiogenic and Morphine-Induced Locomotor Myperactivity.
AD-P003 253	The Defense Nuclear Agency Intermediate Dose Program: An Overview (Effects of Total-Body Irradiation on the Performance of Personnel in Army Combat Crews).
AD-P003 254	A Questionnaire Assessment of Estimated Radiation Effects upon Military Task Performance.
AD-P003 255	Estimated Effects of Ionizing Radiation upon Military Task Performance: Individual Combat Crewmember Assessment.
AD-P003 256	A Summary of the Psychological Effects of Tactical Nuclear Warfare.
AD-P003 257	Physical Performance Tests as Predictors of Task Performance.
AD-P003 258	Affective Determinants of Job Perceptions.
AD-P003 259	Expectations, Job Perceptions and Disconfirmation Among Air Force Academy Graduates.
AD-P003 260	The Effects of REFORGER (Return of Forces to Germany) Exercise Participation on Soldiers' Attitudes.
AD-P003 261	Changes in Attitudes Toward Women at the Air Force Academy.
AD-P003 262	Risk-Taking Performance of Military Personnel: Sex Differences and Practice Effects.
AD-P003 263	The Well-Being of Army Soldiers and Their Families: Methodological, Substantive, and Technical Considerations.

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

(TITLE): Proceedings of the Symposium: Psychology in the Department of Defense (9th) Held at Colorado Springs, Colorado on 18-20 April 1984.

(SOURCE): Air Force Academy, Colorado Springs, CO.

TO ORDER THE COMPLETE COMPILATION REPORT USE AD-A141 043.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDINGS, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 264	Comparing Well-Being and Stress of Several High Risk Groups.
AD-P003 265	Well Being and Distress Among Drill Sergeants: Civilian and Military Comparisons and the Role of Social Support.
AD-P003 266	Habitability and Human Productivity Issues Confronting the Air Force Space Command.
AD-P003 267	Motion Simulation with a G-Seat System: Sensory and Performance Mechanisms.
AD-P003 268	The Measurement of C-5A Aircrew Performance.
AD-P003 269	Aircrew Performance with Simulated Advanced Radar and FLIR (Forward Looking Infrared) Sensors in Single- and Two-Place Crew Stations.
AD-P003 270	Health Risks Associated with Aircraft Model Type Among U.S. Navy Pilots.
AD-P003 271	Initial Validation of a Functional Age Profile for Aviators.
AD-P003 272	A Computer Simulation to Predict Target Identification Latencies.
AD-P003 273	The Design of Manually Operated Controls for a Six-Degree-of-Freedom Groundborne 'Walking' Vehicle: Control Strategies and Sterotypes.
AD-P003 274	Validation of the General Classification Test for Selection of Commissioning from the Ranks Candidates.
AD-P003 275	Validity Generalization of Navy Selector Composites.
AD-P003 276	Improving Navigator Selection.
AD-P003 277	Monetary Value of Navigator Selection Using the AFOOT (Air Force Officer Qualifying Test).
AD-P003 278	Current Research in Royal Navy Officer Selection.
AD-P003 279	Air Force Academy Measures of Performance as Predictors of Promotion Potential.
AD-P003 280	The Effectiveness of Biographical Inventories in Predicting Success in OTS.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 281	USAF-Canadian Forces Exchange of Occupational Analysis.
AD-P003 282	Personal and Organizational Factors Affecting Health and Morale of US Army Special Forces Soldiers.
AD-P003 283	Workshops for Management of Combat Stress.
AD-P003 284	Combat Psychiatry: A Primary Prevention Program in the 4th Infantry Division (Mechanized).
AD-P003 285	Identification Measurement and Prediction of a Psychotic Variable with Neuropsychological Measures.
AD-P003 286	Osychological Evaluation of Aircrew Members.
AD-P003 287	The Relationship between Substance Use, Dermography and Recial/Ethnic Attitudes Related to Performance and Social Attraction.
AD-P003 288	Attitudes Toward Making a Transfer: A Predictive Model.
AD-P003 289	A Path Analytic Study of Psychological Success Among Air Force Officers.
AD-P003 290	Visual Display Technology: A Developing Challenge for the Behavioral Scientist.
AD-P003 291	Animal Behavioral Testing in Medical Defense against Chemical Agents.
AD-P003 292	Animal Models of Human Performance: Structural and Functional Approaches to Extrapolating from Animal to Man.
AD-P003 293	Animal Models for the Assessment of Laser Induced Eye Damage.
AD-P003 294	The Assessment of Nonionizing Radiation Hazards.
AD-P003 295	A Review and Evaluation of Research Concerning the Performance Effects of Nuclear Radiation.
AD-P003 296	Effects of Stress on Individual Productivity, Absenteeism, and Wellness.
AD-P003 297	Stress and Military Women: The Relationship of Job and Life Experiences to Menstrual Distress.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 298	A Review of Open Literature and Recent DoD-Funded Research on Stress in Organizations.
AD-P003 299	Aftereffects Associated with One and Two Stressors Across Conditions of Complete, Partial, or No Control.
AD-P003 300	Psychological and Biochemical Effects of a Stress Management Program.
AD-P003 301	Stress Management in Israeli Defense Force: Implications for U.S. Military Human Service Professionals.
AD-P003 302	Military Leadership Education: What War College Journals Seem to Suggest.
AD-P003 303	The Role of Relevant Experience and Intellectual Ability in Determining the Performance of Military Leaders: A Contingency Model Explanation.
AD-P003 304	There is No Such Thing as Leadership.
AD-P003 305	Productivity, an Organizational Rather Than an Individual Variable.
AD-P003 306	Leadership Instruction and Officer Candidates: Some Unholy Thoughts.
AD-P003 307	Windows on a Future World: Human Dimensions of Air Force Doctrine in the 21st, Century.
AD-P003 308	Performance Based Management Systems: Filling Army Human Resource Demands in the Future.
AD-P003 309	Computer-Automated Technological Innovation in Three Manufacturing Sectors.
AD-P003 310	Prototype Development of an Information-Sharing and Decision Support System for the Manpower Personnel and Training Community.
AD-P003 311	An Identification of Actual Versus Reported Accident Causes.
AD-P003 312	Relationships of Biorhythms and Catergories of Industrial Accidents on an Air Force Installation.
AD-P003 313	Management Information Systems: A Need for Human Factors.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 314	Individualizing with Computer-Assisted Instruction (CAI): Enhancing Student Motivation through Computer-Controlled Audio.
AD-P003 315	Training and Performance Support System: A New Dimension in Individualized instruction.
AD-P003 316	Maintenance Troubleshooting Simulation: Adapting to the Trainee.
AD-P003 317	Evaluation of the Electronic Equipment Maintenance Trainer (EEMT): Individualized Troubleshooting Training.
AD-P003 318	Factors Influencing the Selection of Individualized Instructional Approaches.
AD-P003 319	The Structure of Processing Resource Demends in Monitoring Automatic Systems.
AD-P003 320	Magic (Microcomputer Applications of Graphics and Interactive Communications) 1: A Study of Manual Versus Vocal Control under Heavy Task Loading.
AD-P003 321	The Limits of Multiple Resource Theory in Display Formatting: Effects of Task Integration.
AD-P003 322	Tracking with Intermittent Radar Coverage: I. Interruptions After Each Collected Frame of Imagery.
AD-P003 323	Tracking with Intermittent Radar Coverage: II. Interruptions After Two or More Consecutively-Collected Frames of Imagery.
AD-P003 324	Family and Work in the Air Force.
AD-P003 325	The Changing Roles of Military Wives: An Analysis of Work-Related Values and Lifestyle Variables.
AD-P003 326	Isolation Matrix: A Tool for Discovery.
AD-P003 327	A Pilot Test of the Use of Telephone Methods to Survey Active Army Personnel.
AD-P003 328	Linkages between Family Support Variables and Military Career Commitments.
AD-P003 329	Perceptual Factors in Low Altitudes Flight.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 330	Political Ideology and the Willingness to Enlist.
AD-P003 331	Young Germans between Compulsory Military Service and Conscientious Objection.
AD-P003 332	The Impact of Information Quality on Career Decision Making.
AF-P003 333	Interelations Among Career Counseling, Organizational Commitment and Career Satisfaction.
AD-P003 334	Identifying Variables Influential to an Officer Promotion Nomination Board.
AD-P003 335	The Role of Retention in Managing the Aviation Warrant Officer Force.
AD-P003 336	The Relationship of Role Congruity to Air Force Pilot Turnover and Performance.
AD-P003 337	Psychophysiological Tools in Engineering Psychology.
AD-P003 338	Individual Differences in Learning Rate.
AD-P003 339	Two for the Money--Speed and Level Scores from a Computerized Vocabulary Test.
AD-P003 340	Assessment of Learning Abilities Using Rate Measures.
AD-P003 341	Characteristics of Navy Recruits on Reading Comprehension and Related Variables: A Quarterly Three-Year Profile for FY 1981-83.
AD-P003 342	MMPI (Minnesota Multiphasic Personality Inventory) Normative Data for a Male Active Duty Army Population.
AD-P003 343	Age Effects on Active Duty Army MMPI (Minnesota Multiphasic Personality Inventory) Profiles.
AD-P003 344	Enlisted Performance Evaluation System for the U.S. Coast Guard.
AD-P003 345	Use of Self Assessment in Estimating Levels of Skill Retention.
AF-P003 346	Comparing Two AFOOT (Air Force Officer Qualifying Test) Norming Groups Using Weighted Multidimensional Scaling.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 347	Training Decision-Makers to be Creative: A Management Process Model.
AD-P003 348	On the Meaning of the 'Whole Person' Concept and Achieving Wholeness.
AD-P003 349	Behavioral Issues in the Management of Technology.
AD-P003 350	Methodological Issues in the Study of Air Force Organizational Structures.
AD-P003 351	Leadership, Managership, and Computers in Today's Air Force.
AD-P003 352	Leadership on the Future Battlefield: Systems-Wide Perspectives.
AD-P003 353	A Systematic Comparative Analysis of Conditions Analogous to Long-Duration Space Missions.
AD-P003 354	Criteria for Selecting Subjects for the Assessment of Advanced Crew System Concepts.
AD-P003 355	Relationship between Crewmenber Characteristics and Tank Crew Performance.
AD-P003 356	Field-Expedient Maintenance Experiences of M60-Series Tank Crewmen.
AD-P003 357	Goals and Feedback in Command Group Training.
AD-P003 358	An Examination of Air Force Supervisor-Subordinate Communication Relationships and Outcome Variables: A Review of the Findings.
AD-P003 359	Terraih Travel Simulation: Data and Application.
AD-P003 360	Static versus Dynamic Presentation of Visual Cues in Simulated Low Level Flight.
AD-P003 361	NOSC (Naval Ocean Systems Center)-Hawaii Perceptual Sciences Research Program.
AD-P003 362	Factors Affecting Visual Performance with Steroscopic Television Displays.
AD-P003 363	Interactions of Motion Perception with Pattern Perception.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P003 364	Using Comparison-Based Methods for Predicting and Designing.
AD-P003 365	The Dual Meaning of the Term Context: Implications for Research, Theory and Applications.
AD-P003 366	Research Integration: An Essential for Department of Defense Psychological Research.
AD-P003 367	What Training is Depends on Who You Ask.
AD-P003 368	Video Games: A Human Factors Guide to Visual Display Design and Instructional System Design.

PANEL SESSION

PANEL SESSION: TESTING THE HUMAN DIMENSIONS OF COMMAND, CONTROL, COMMUNICATIONS AND INTELLIGENCE (C^3I) SYSTEMS

SESSION CHAIR: Elizabeth J. Kirk (The MITRE Corporation)

PANELISTS: Elizabeth J. Kirk (The MITRE Corporation)

Dennis K. Leedom (Office of the Deputy Under Secretary of Defense)

Morton Barov (The MITRE Corporation)

PROCEEDINGS ENTRIES

"Environments for evaluating the performance of C^3I systems"
(Elizabeth J. Kirk)

"A sociotechnical communications model for command and control"
(Dennis K. Leedom)

Evaluation of C^3 systems using command post exercises"
(Morton Barov)

Environments for Evaluating Performance of C^3I Systems

Dr. Elizabeth J. Kirk, The MITRE Corporation

Abstract

Some environments and methods which have been used to evaluate C^3I systems are outlined. A series of specific controlled tests that were conducted during recent (1982-83) developmental tests of Joint Service message standards are also described. This type of test may be seen as a cost-effective method to evaluate certain system capabilities in addition to or in place of large-scale field tests.

I. Introduction

While the test and evaluation processes that must take place during the development and acquisition of materiel, including command, control, communications, and intelligence (C^3I) systems, are clearly outlined in military documents (e.g., Army regulation 70-10, 71-3), the environments in which testing takes place, and the testing methods to be used are often left to the discretion of developers and military testing agencies. Often system tests are piggy-backed onto operational exercises rather than designed to meet the specific needs of the system function being tested. The purpose of this paper is twofold:

- o To familiarize symposium participants with some of the more common Test and Evaluation (T&E) environments and methods that have been employed during the system acquisition cycle. Other papers presented on the T&E panel will use one or more of these environments or methods.
- o To report on some controlled tests that were conducted as part of Army developmental testing of Joint Service message standards. The purpose of the controlled tests was to determine the difference between various message preparation modes and procedures.

II. Test and Evaluation

A. The Acquisition Cycle

The current Army Materiel Acquisition process includes four phases: conceptual, validation, full-scale development, and production with testing taking place during the last three phases in the form of Developmental Tests (DT) and Operational Tests (OT). According to Army regulations, DT is conducted "to demonstrate that the engineering design and development process is complete, that design risks have been minimized, and that the system will meet specifications; and to estimate the system's military utility when it is introduced." The tests are planned, conducted, and monitored by the developer.

OT, on the other hand, is conducted to estimate a prospective system's military utility, operational effectiveness, operational suitability, and the need for modifications. OT also provides information on organizational, manpower, doctrinal, and tactical requirements. OT tests are executed by personnel who would be expected to use and support the system when deployed. OT tests are to be conducted in "as realistic an operational environment as possible."

C^3I systems present challenges to the T&E process for systems acquisition which was originally designed for weapons systems and other materiel. The proliferation of command nets between and among services and the use of computers in command centers pose very new and complex problems for evaluating system functional capabilities relating to compatibility, interoperability, software design, and man-

machine interfaces. With many of these systems, the human becomes an integral part of overall system performance. Not only must individual human performance capabilities be ascertained in these areas, but small group or team performance, and even organizational performance must be tested. The means to find suitable ways to identify areas to be tested, the environments to test them in, and test designs and methods are not clearly defined, leaving both developers and testers to generate their own T&E procedures. The new demands placed on T&E because of the complexity of systems raises two very important issues:

- o What future tests can be designed and conducted to cost-effectively evaluate a system's capability to perform certain functions at the individual, team, and organizational levels? Answering this question takes much forethought and planning since many types of test methods may have to be used to test thoroughly all of the system's capabilities at each stage of development and at each level.
- o What can be done to use the current set of test environments and cycles more fruitfully? Can a series of system tests, using various methods, be imposed on a Command Post Exercise (CPX), for example?

B. T&E Methods

CPXs, FTXs, laboratories, proving grounds and factories where the C^3I system is being developed can all provide test environments. Each provides its own set of strengths and weaknesses in terms of control over test conditions, the pool of "subjects," data collection capabilities, and cost. Some or all of the methods outlined below can be used in such environments.

1. Case Studies. Many system T&E reports are based on performance during a particular exercise (i.e., how system X performed during Solid Shield '83). Recommendations for changes are made on the basis of these reports and, if possible, checks on improvements made are followed up in future tests. However, these checks are often overlooked because the lessons learned from one previous exercise are not carried over to the next. New testers may not read old reports; exercise goals can change test conditions to the point that it is impossible to replicate previous tests. Thus, case study evaluations based on one test or exercise may not provide the continuity, reliability, and replicability to thoroughly exercise all system functions and evaluate its capabilities.

Efforts can be made to assure that the information collected during one exercise is consistent (that is, the same measures were used and the same questions asked across various users, elements, and time periods) and that the same types of data are collected over a series of exercises. Efforts also should be made to ensure that one is able to state which capabilities were affected by the test environment and which by system performance.

2. Controlled tests. When specific system functions must be evaluated and require either many repetitions of tasks or repititions under strictly controlled conditions, then controlled tests can be conducted. The system being tested must be in some way isolatable from the rest of the systems or personnel by either controlling for (statistically or in the test design) or holding constant those influences not being examined for the specific function under evaluation. Controlled tests can include a few individuals using one system or a whole network of systems and personnel performing assigned functions.

3. Expert review. Teams of "experts" (operational or technical) can be used at the early stages of system development to define and refine functional requirements. They can also be used at later stages to evaluate a system's ability to meet engineering design and operational requirements. For example, one "expert" might say: "Because of the introduction of system X in the field, the user of system Y must now also receive, store, retrieve, and display information collected from system X.

Does it have this capability?" In this way, a system's functional capabilities can be evaluated in the operational realm, in part, without conducting large scale exercises and tests.

4. Modeling and Simulation. Modeling system capabilities before the actual development of a prototype can alert developers to problems within the system. This is especially useful during developmental testing in the early stage of acquisition. Simulation can also be used to simulate the larger environment into which a system is to be placed. In this case, simulation can provide a cost-effective approach to testing system capabilities in a larger context without requiring all the manpower and equipment needed for a large-scale exercise.

In all stages of DT and OT test and evaluation, the following issues should be considered.

1. What functions of the system must be tested? Which functions involve a one-human-one-machine, small group, or large organizational context?

2. What operational exercises or tests currently exist to evaluate machine and human components of system effectiveness? How can these operational exercises or tests be exploited to collect good data efficiently?

3. What additional tests must be planned to assure complete test and evaluation of the system.

While these questions may seem obvious, the planning of tests in conjunction with the develoment timetable becomes very complex when developer, tester, and operational forces exercise timetable must be coordinated.

III. The Use of Controlled Tests to Assess Message Preparation Processes.

This final section briefly describes a series of controlled tests that were conducted as part of a set of larger developmental tests. While message standards tests may not be the same as other system tests, the point I want to make here is the usefulness of controlled tests that have been piggybacked on top of other tests which were not designed to answer some very specific questions. This type of method can be easily adapted to a large-scale test with careful planning and sponsor suport.

The larger test, as part of the Joint Interoperability of Tactical Command and Control Systems (JINTACCS) program, were conducted to determine the compatibility and interoperability of standardized messages exchanged between services. The JINTACCS program objective is to establish and test tactical C^2 data exchange standards that are adequate for joint service operations and usable for both current manual interfaces and future automated interfaces. Tests must be conducted to determine if:

- o standard formats can be established that are readable and writeable and are usable for data transmission;
- o the standardized messages contain sufficient information to conduct military operations; and
- o standards can be established for use in man-man, man-machine, and machine-machine interfaces.

A series of joint developmental tests was conducted in which all of the services involved in intelligence, air, operations control, and fire support operations exchanged messages. In addition to the jointly sponsored tests, the Army's Test Management Division of the Center for Systems Engineering Integration (CENSEI) supported a series of controlled tests, oftentimes run concurrently with the Joint Tests, to answer issues germane to Army-specific interests. The results of these controlled tests are quite preliminary and more testing must be done to confirm the findings. However, the utility of using such controlled tests to answer specific questions is supported by the findings.

For each of the tests a test design was prepared. Each design attempted to control for differences between subject, stimuli, and schedule of events. In the cases

where the number of subjects was insufficient to provide definitive results using statistical techniques, the results provided evidence for suggesting system or standards improvement. Three tests conducted yielding the most conclusive results are outlined below:

A. Tactical Computer Terminals (TCT) Tests (Turner, et al, 1982)

1. Test Question: Can messages be prepared as rapidly and accurately using the TCT as they can manually?

2. Purpose: To compare manual preparation processes with automated (TCT) preparation processes.

3. Background: The TCT is a militarized computer terminal with 96 K of memory. The Army is considering it as an interim system to be used for automated JINTACCS message preparation. JINTACCS message formats, codes, and preparation instructions are coded in the software and can be used to directly prepare and transmit JINTACCS messages. In the test, four military personnel were asked to prepare two messages each using both the manual and automated (i.e., TCT) methods of preparation. Message preparation times and errors and participant comments were recorded.

4. Results:

- Manual and automated (TCT) preparation and transmission took about the same time.
- The subjects made about the same number of errors using either mode.
- Software and editing changes on the TCT could improve TCT message preparation time and accuracy. Recommendations were made.

5. Discussion: This kind of test provided valuable insight into the kinds of software changes that are desirable to automate JINTACCS standards and message preparation. Such tests did not require that automated systems be plugged into a larger network, although for questions of compatibility and interoperability with other Army and Service systems, this would be desirable at some later date.

B. Automated Reference Test (Kirk, Plotkin, 1983)

1. Test Question: Can the material in the Army JINTACCS User Handbook be accessed as fast or faster using an automated system as the hard copy Handbook?

2. Purpose: To compare standard hard copy access rates to an automated access rate.

3. Background: The Electronic Media System (EMS) consists of an Apple II computer, a videodisk, a videodisk player, and a monitor. It was used as a computer-aided access to look up material in an automated version of the Army JINTACCS User Handbook. This handbook provides the soldier with standard message format, instructions for their use, and, when needed, standardized entry codes for pieces of information. The handbook was automated so that the user could quickly "flip through" pages of the handbook. Two civilian personnel were given a series of tests to access different sections of the User Handbook using both the EMS and hard copy material. Access times were recorded.

4. Results:

- Access to the manual (hard copy) and automated (EMS) materials was equally fast.
- Software improvements on the EMS could possibly improve access speeds.
- The EMS could have broader applications as a JINTACCS training or briefing device.

5. Discussion: This test indicated that, at least in this application, automation is not always "better" than manual methods. Further examination of the EMS concluded that its strengths could be exploited for other purposes in the JINTACCS program.

C. Standardized Message Format Test (Kirk et al, 1983)

1. Test Question: Can messages be prepared as rapidly and accurately using blank forms as pre-printed structured ones?

2. Purpose: To determine the necessity of pre-printed standardized forms for preparing short messages.

3. Background: All JINTACCS messages currently prepared by the Army use pre-printed forms. The current procedure requires user elements to maintain a constant supply of these forms. Two soldiers were asked to prepare a series of Fire Support messages using both the pre-printed and blank forms. Timing and error data were recorded.

4. Results:

- Blank forms and pre-printed message forms took the same amount of time to prepare.
- There were no significant differences in the number of errors committed using the two forms.
- For short, linear messages, it may be possible to eliminate pre-printed forms. More tests should be run on all message types to verify these findings.

5. Discussion: This test raised the larger issue of the "paper flow" problem in preparing large numbers of structured written messages and distributing them to appropriate destinations in a timely fashion. These preliminary results suggest that some of the problem can be alleviated. However, the logistics of assuring that proper forms are distributed effectively are subject to test in a larger exercise environment.

IV. Conclusions

The current systems acquisition process allows for a number of developmental and operational tests to be conducted throughout all stages of development. C^3I systems place new demands and challenges for test and evaluation of the machines and men operating them. While case study and field exercises have been the method and environment used extensively to test systems, a wide range of methods and environments exist to test and evaluate systems during all phases. The method and environment should fit the particular function of the system being tested rather than the test being fit into a particular T&E environment that is available.

As described above, controlled tests can provide a fertile yet relatively inexpensive environment to test selected aspects of system and user performance. They can also be included as part of a larger exercise with careful forethought and planning and sponsor support.

References

Kirk, E. J. and N. J. Plotkin; Test Report on the Electronic Media System (EMS Acceptance Test; McLean, VA; The MITRE Corporation; WP82W00675; 1 April 1983.

Kirk, E. J., N. J. Plotkin, L. S. Russell, and F. J. Stech; Test Report on JINTACCS Fire Support Developmental Certification Testing; McLean, VA; The MITRE Corporation; WP83W00472; 15 September 1983.

Turner, T., E. J. Kirk, and N. J. Plotkin; Test Report for Air Operations Compatibility and Interoperability Tests 1 and 2; McLean, VA; The MITRE Corporation; 21 December 1982.

A Sociotechnical Communications Model for Command and Control

Dennis K. Leedom
Office of the Deputy Under Secretary of Defense
(Communications, Command, Control and Intelligence)

Abstract

Using military command and control systems as the setting, this paper examines the problem of effectively communicating perceptions and mental concepts from one organizational element to another. A model is proposed which identifies and links the important aspects of this process: problem formulation, language development and usage, and transmission media. Using this model, the paper outlines tradeoffs among these various factors and suggests ways in which command and control processes might be improved through both technical and non-technical means.

Introduction

An essential ingredient for any organization is the development of effective communications among its various components. Often, as in the case of distributed command and control systems, our concern focuses on the technical means of transmitting voice or data messages from one point to another. Yet, as pointed out by the Defense Science Board (1978) and the Armed Forces Communications and Electronics Association (1982), technology is not the limiting factor in developing effective command and control systems. In contrast, we must begin to look not only at how we communicate, but also at what we communicate. This latter issue, however, moves us beyond questions of technology and into the cognitive processes of the military decisionmakers --into the culture and language of combat.

Several aspects of combat operations suggests the need to examine the entire process of how perceptions and mental concepts are communicated among different staffs within a command and control system. First, battlestaffs face a number of real-world considerations which preclude a precise, structured approach to the planning and execution of military operations. Hence, a significant portion of the command and control process involves definition of the combat problems to be solved and the decisions to be made. Second, these staffs do not generally deal directly with the battlefield; rather, they deal with perceptions of the battlefield which have been developed through the eyes and ears of subordinates. If a distributed command and control system is to respond in a coordinated and consistent manner, then each of its constituent parts must share a common perception of the battlefield, the combat operations which unfold upon the battlefield, and the strategies and tactics which shape these operations.

In the following paper, I develop a conceptual model for addressing the problem of shared perceptions within a command and control system. By

necessity, such a model must consider the cognitive aspects of expert knowledge and problem solving, the psycholinguistic aspects of communicating ideas from one culture to another, and the sociotechnical aspects of transmitting these communications.

A Conceptual Model of Organizational Communication

A sociotechnical model for examining organizational communications is outlined in Figure 1. It is through such a model that we can relate the various cognitive, psycholinguistic and technological factors which influence the nature and effectiveness of such communications.

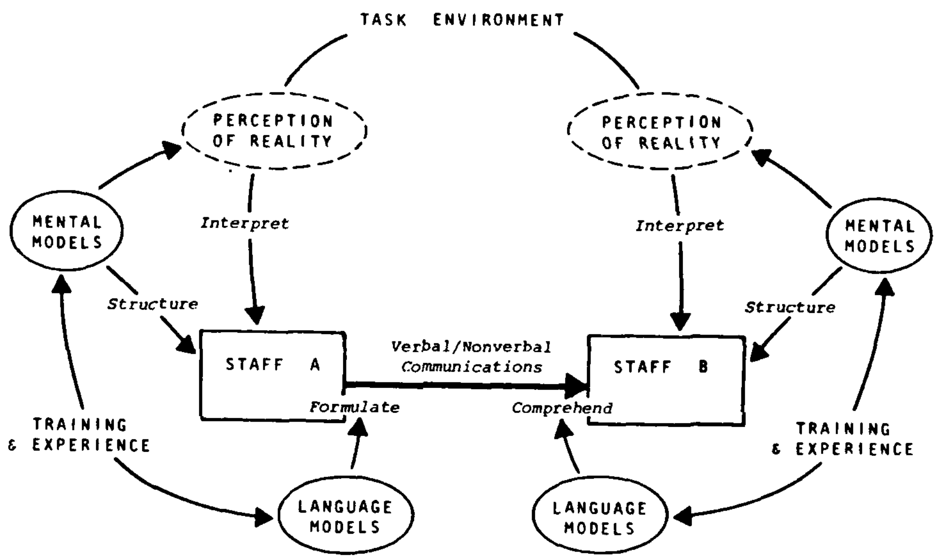


Figure 1 A Sociotechnical Model of Organizational Communications

Imagine that two staffs must develop a shared perception of some common task environment in order to effectively coordinate their operations. These staffs might represent Army and Air Force target analysts located at two different headquarters. Their common task might be the development of battlefield interdiction targets to be struck by either long-range surface-to-surface missiles or tactical fighter-bomber aircraft.

The communication process begins with "Staff A" sampling the task environment, extracting salient cues or other information, and codifying this information into a meaningful perception of the situation. As emphasized by Weick (1969), this process serves to resolve environmental uncertainty (or equivocality) --to establish a workable level of informational certainty. For our interdiction example, the structuring and interpretation of battlefield reports into meaningful patterns and events are essential steps in the operational planning process.

As indicated in the figure, the structuring and interpretation of information is guided by the application of mental models which have been developed through past training and experience. Here, both generative and

interpretive processes are postulated to occur. Examples of generative processes are described by Simon and Chase's (1973) concept of information chunking and Druzhinin and Kontorov's (1972) concept of axiomatic thought. In the former case, we use "chunks" of meaningful information to overcome our apparent limitations in short-term memory and comprehension. In the latter case, axiomatic (or rule-based) thought permits us to reach rather sophisticated levels of reasoning through the hierarchical application of simple rules or relationships.

Examples of interpretive processes are illustrated by Minsky's (1975) theory of frames and scripts, Simon and Hayes' (1976 and 1977) model of isomorphic problem solving, and Gick and Holyoak's (1980) concept of analogical reasoning. In each case, we organize our observations from the real world according to a set of standardized expectations and beliefs. These more holistic mental models assist us in focusing our attention and efficiently processing informational inputs; but, at the same time, they can dangerously mislead us if we erroneously adopt stereotypical beliefs about an actual situation.

An important aspect of mental models is that they are the products of our unique learning experiences. That is, our interpretation of events and situations surrounding us is largely shaped by our individual knowledge, experiences and positions of responsibility. Given the loosely-coupled nature of decisionmaking processes within a command and control system, it is important for each staff to employ consistent mental models in developing their interpretations and responses. For the interdiction example, this consistency would be reflected in a cohesive and well-coordinated attack against the enemy's second-echelon forces. Such consistency may be achieved in several ways: through the development of common doctrine and tactics, through prior training and joint exercises, through the use of jointly developed decision aids, or through effective communications. However, each approach depends upon the language we use to shape and express our ideas.

From Vygotsky (1934), we know that "thought is not merely expressed in words; it comes into existence through them." This notion is expressed more formally in the Sapir-Whorf hypothesis (Sapir,1929 & Whorf,1956) which holds that we interpret reality through the particular language used by our cultural group. Within organizations, language differences are seen to emerge whenever one subunit develops an area of operational expertise and, hence, moves beyond other subunits in its conceptualization of a problem (Hall,1982). In contrast with this notion, Chomsky (1965) argues that all languages are comparable since they share common principles of grammar and syntax. Combining these two points of view, we see the importance of developing common languages. That is, common languages permit the efficient communication of understandings across cultural divisions.

Such divisions might exist between Army and Air Force battlestaffs which perceive and speak of the battlefield according to different terms, concepts and interests. Lacking a common language of combat, battlestaffs from each Service might be required to engage in considerable dialogue or message exchange in order to clarify target priorities, attack strategies and so forth. As will be seen in a moment, communication efficiency has significant implications for the types of message transmission systems required for command and control systems.

The connection between language and thought is further reflected in the general trend of current psycholinguistic research. As reported by Danks and Glucksberg (1980), researchers have shifted their focus from a word-based, associative model of language comprehension to a knowledge-driven, cognitive model. Hence, many of the same semantic constructs used for codifying expert knowledge are now seen as the mechanisms underlying language comprehesion. Fisher (1981), in his analysis of organizational communications, provides further support for this model in defining typical misuses of language. From such misuses, we know that communication failures result when we disrupt the recipient's process of interpreting our message within a familiar conceptual framework.

So far, we have seen that communication processes are inextricably related to the mental models held by the sender and recipient. Further, we note that considerable dialogue is often required for achieving comprehension when the communicators do not share a common language. It is at this point where we must address the technical means by which we communicate.

Linkage of the cognitive and technical aspects of communications is provided through Daft and Lengel's (1984) concept of information richness. Within this concept, information richness is defined as the potential information carrying capacity of data. In turn, the richness of a particular transmission media is defined by the availability of immediate feedback, the number of verbal or nonverbal communications channels used, the source of the communication and the type of the language used. Face-to-face communications provide a high richness media while computer-generated numerical data reports represent a low richness media.

The significance of Daft and Lengel's work lies in the requirement to match the richness of the communications media to the complexity of the supported decision process. Within most organizations, decision complexity is seen to be a function of subunit differentiation (differing frames of reference) and subunit interdependence (level of information exchange required). A mismatch can occur in either of two directions: overcomplication (too many cues, noise, ambiguity) or oversimplification (too few cues, impersonal, no feedback).

Sociotechnical Improvements for Command and Control

In applying the concept of information richness to military command and control systems, it is important to note that practical limits may be imposed on the level of richness available. These limitations might arise due to enemy countermeasures, physical separation of command posts or the exigencies of combat. As a result, we may face situations in which the available richness of the communications media does not match that level required by the complexity of the supported decision process. It is at this point where the sociotechnical model suggests a number of improvements.

As the name sociotechnical implies, we must look at both technical and non-technical options for improving our command and control capabilities. Technical options generally fall under the category of improved communications media (eg, higher bandwidth transmission channels, higher

quality secure voice and anti-jam media). However, a number of non-technical options also exist which have generally received less attention:

> Joint doctrine development: building agreement over combat objectives, priorities and basic strategies --all essential ingredients to the mental models used by each battlestaff
>
> *Joint training schools*: *refinement of joint doctrine into* mutually understood concepts of operation and language of combat
>
> Joint training exercises: further refinement of concepts and development of efficient languages for communication
>
> Jointly-developed decision aides: development of "corporate memory" and standard cognitive constructs.

Through the above options, it will be possible to develop a common set of mental models and operational language. Together with our technical advances in communications technology, these options represent a total-systems approach to improving our command and control systems.

Summary

In summary, the sociotechnical model of communications provides us with a broad conceptual framework for addressing military command and control systems. The approach is interdisciplinary and, as a result, this increases resistance to its adoption. For the future, such resistance should be overcome in our research and development efforts as we better understand both the technical and social nature of command and control systems.

References

Chomsky, N., Aspects of the Theory of Syntax (Cambridge, Mass: MIT Press, 1965)

Command and Control (C^2) System Acquisition Study Final Briefing (Washington, D.C.: The Armed Forces Communications and Electronics Association, July 1982)

Daft, R.L. and Lengel, R.H., "Information Richness: A New Approach to Managerial Behavior and Organization Design," to appear in Research in Organizational Behavior (JAI Press, 1984, Vol 6, Staw, B. and Cummings, L.L. eds)

Danks, J.H. and Glucksberg, S., "Experimental Psycholinguistics," Annual Review of Psychology (1980, Vol 31, pp 391-417)

Druzhinin, V.V. and Kontorov, D.S., Concept, Algorithm, Decision (A Soviet View) (Moscow, 1972) (trans: United States Air Force, 1976)

Fisher, D., Communication in Organizations (St. Paul, Minn: West, 1981)

Gick, M.L. and Holyoak, K.J., "Analogical Problem Solving," Cognitive Psychology (Academic Press, 1980, Vol 12, pp 306-355)

Hall, R.H., Organizations, Structure and Process (Englewood Cliffs, N.J.: Prentice-Hall, 1982) (3rd ed)

Minsky, M., "A Framework for Representing Knowledge," The Psychology of Computer Vision by P.H. Winston (New York: McGraw-Hill, 1975)

Report of the Defense Science Board Task Force on Command and Control Systems Management (Washington, DC: Office of the Under Secretary of Defense for Research and Engineering, July 1978)

Sapir, E., "The Status of Linguistics as a Science," Language (1929, Vol 5, pp 207-214)

Simon, H.A. and Chase, W.G., "Perceptions in Chess," Cognitive Psychology (Academic Press, 1973, Vol 4, pp 55-81)

Simon, H.A. and Hayes, J.R., "The Understanding Process: Problem Isomorphs," Cognitive Psychology (Academic Press, 1976, Vol 8, pp 165-190)

Simon, H.A. and Hayes, J.R., "Psychological Differences Among Problem Isomorphs," Cognitive Theory (Castellan, N.J.: Lawrence Erlbaum Assoc, 1977, Vol 2, pp 21-41, Pisoni, D.B. and Potts, G.R. eds)

Vygotsky, L.S., Thought and Language (Cambridge, Mass: MIT Press, trans 1962) (first published in USSR, 1934)

Weick, K.E., The Social Psychology of Organizing (Reading, Mass: Addison-Wesley, 1969) (2nd ed. 1979)

Whorf, B.L., "Science and Linguistics," Language, Thought and Reality (Cambridge, Mass: MIT Press, 1956, pp 212-213, Barroll, J.B. ed)

EVALUATION OF C3 SYSTEMS USING COMMAND POST EXERCISES

Morton Barov
The MITRE Corporation

Evaluation of command and control (C2) systems is a topic of increasing concern within the defense community. Particularly when systems are acquired using an evolutionary rather than traditional development approach, methods are needed to gain meaningful feedback from C2 system users early in the development process.

This paper discusses the use of the Command Post Exercise (CPX) environment to generate user input to the evolutionary development of command and control systems. This evaluation approach was developed and implemented over a five-year period based on experience with Army participation in CPXs in Europe. In the approach, prototype command and control subsystems were placed in the hands of the system users during regularly scheduled CPXs; user experiences with prototype system capabilities were used to guide evolutionary development.

The key points made in this paper concerning the evaluation of C2 systems can be summarized as follows:

- User involvement at all stages of system development-- from the definition of objectives to operational testing of system components-- is key to the evolutionary development of C2 system development and would equally benefit C2 system development which follows a more traditional approach.

- Command Post Exercises (CPXs) provide a valuable opportunity to obtain needed user feedback to C2 system development.

- Successful use of CPXs as an evaluation tool depends upon substantial front-end planning, both to ascertain what aspects of systems can reasonably be evaluated in a given CPX and to develop applicable methods of data collection and analysis.

- Ideally use of CPXs would be implemented as one part of a systems approach to test and evaluation in which a mix of evaluation approaches would be employed to address changing evaluation needs through the development process.

- Finally, evaluation of C2 systems, if it is to contribute to the development of effective C2 capabilities, must address issues of both system performance and system utility. While it is necessary to have systems which perform in accordance with stated performance requirements to be sufficient, C2 systems must support real user needs.

AD-P003 239

Military Deception: A Clinical Psychological Analysis

Richard W. Bloom
Armed Forces Medical Intelligence Center
Ft. Detrick, Frederick, MD 21701

Abstract

This paper describes four clinical psychological contributions to the study and practice of military deception. First, inferential techniques yield five generic factors of deception: intention, behavior, perception, teleology, and systems. Second, reliability studies and causal modeling techniques applied to these factors are slowly leading to the development of a prescriptive, deception data base. Third, characteristics of those who deceive well versus those susceptible to deception are being identified through modified task analyses. Fourth, inferences are being made as to the reliability, validity, and deception potential of foreign research in behavioral sciences. The above have relevance for the selection, training, and performance of military deception planners and analysts.

Military deception. Sun Tzu (400 B.C.) terms it the supreme martial art, a means to the greatest gain with the least effort. From Thucydides (350 B.C.) we infer it as the optimal achievement of warfare, as mind over mind instead of mind over matter. On the other hand, Clausewitz (1831) downplays military deception, probably because it violates the heroic ethic and military virtues.

In our era we can at least agree that military deception has worked and has worked well. Examples abound from the Chinese Civil War through Operation Peace for Galilee, from the turmoil of battle to detente, from the tactical to the strategic (e.g., Stuart & Tow, 1982).

However, there have been few attempts to systematize what deception is and how it works. Classics such as the Chan-kuo Ts'e (50 A.D.) and the Muqaddimah (1370) are merely anecdotal. Modern works comprise either theory without empirical grounding (Daniel & Herbig, 1982) or dust-bowl empiricism with "bean counts" that may include apples and oranges (Whaley, 1969). However, all sources agree that military deception involves the assessment and subsequent influence of a target. This terrain is shared with a host of other pursuits--advertising, sales, and clinical psychology.

This paper describes four contributions to the study and practice of military deception based on theory, experiment, and observation from clinical psychology. While contributions from advertising and sales may be as valuable, they would certainly not be as popular with an audience of psychologists.

1. Generic Factors of Military Deception. What is military deception? I've applied a systematic, inferential technique to historical accounts of military deception from different cultures, societies, conflicts, and eras. Following Levy (1963) I used a three-step semantic stage and a one-step propositional stage. The first semantic step used congruence and dissonance as permissable criteria. The second used structure, function, effect, affect, sequence, frequency, intensity, and pervasiveness as interpretive perspectives. The third used contextual consistency and my own bias for parsimony to support the potential validity of the interpretive perspectives

in step two. The one propositional step used subsumption and understanding as generalization criteria.

This inferential procedure yielded five generic factors: intention, perception, behavior, teleology, and systems.

Intention. The initiator of a deception takes an action or nonaction intended to direct the target away from the truth as the initiator perceives it. In a successful deception, the action or nonaction must be accurately received, analyzed, and disseminated by the target. The intention, on the other hand, must not be.

Perception. The initiator's action or nonaction leads to the target misperceiving reality. This misperception can involve cognition and affect, any psychic process or product.

Behavior. The target's misperception must have a significant influence on target behavior--proximal or distal. In military deception there are at least 108 classes of behavior that can be influenced. These involve the combination of three military areas (operations, intelligence, research and development), three properties of these areas (structure, function, and process), three resource classes (money, materiel, personnel), and four resource properties (time, place, type, quantity/intensity).

Teleology. The target's behavior must have an ultimate end or purpose that satisfies a need of the initiator. This need can be satisfied by directly improving the initiator's lot, worsening that of the target, or improving the initiator's position relative to the target.

Systems. Military deception occurs in a system of initiator, primary target, secondary target, and a context wherein multiple deceptions are ongoing. The secondary target may be an intermediary one, merely an observer of ongoing deception, a future primary target, or a moderator influencing another target otherwise immune to deception. The context consists of politico-military, economic, sociocultural, biomedical, and physical (e.g., geographic) information with local, regional, and global implications.

That military deception includes an initiator, primary target, and secondary target links it with triadic, influence techniques in family therapy (e.g., Haley, 1977).

2. Military Deception Data Base. A data base should cogently describe what has worked and suggest what will work in military deception. Two approaches in these endeavors are the formal and the contextual.

In the former, categories and accompanying Likert rating scales are being developed and standardized for each of the five generic factors. An initial category pool was obtained by the author making up labels for specific instances of the five factors as the latter seemed to appear in deception accounts. Then subjects rated the applicability of each of the category labels for each deception account.

Likert rating means have been in the "applicable" direction for at least some category labels for each factor for each deception account. Split-half and internal measures of reliability for category ratings within factors have been lower than classical psychometric standards dictate as acceptable. Temporal measures of reliability have approached these standards. Factor analyses of the intercorrelational matrices of all category ratings among factors have varied significantly in output. We would hope to extract, of course, some semblance of what was inferred from the initial derivation of the five generic factors of military deception.

These reliability problems have three saving graces. First, even by classical psychometric standards, reliability does not necessarily limit

predictive validity (Karon, 1968). And prediction is the sine qua non of military applications of psychology. Second, although construct validity may be threatened, the five generic factors have proven valuable as planning format aids. Here "valuable" is measured by the self-report of military deceivers (Bloom, in press). Third, the validity of the five generic factors and their associated categories may be supported by structural equation models with latent variables.

This last possibility may be accomplished by assuming the five generic factors are unmeasurable, hypothetical constructs. We then assume that they are related in, let's say, a unidirectional, linear, and sequential manner. Following this, we relate the five generic factors to the list of categories--the latter are measureable. When all relationships are specified in mathematical form, we have a model with a certain structure and unknown parameters. We can then accept or reject the model depending on how well it explains the statistical properties of the measured variables in terms of the hypothesized latent variables.

To my knowledge, this has not yet been accomplished. Difficulties include the sparse documentation of too many deception accounts and the frequent inadequacies of split-half, internal, and temporal reliabilities for the more abstract categories of the systems factor. The latter is also a problem in personality research (Kenrick & Springfield, 1980).

As opposed to the above formal approach, I've also been working with a contextual approach. This involves what Sarbin (1982) terms "emplotment," the attempt to delineate the target's ongoing narrative or script, and what story the initiator must weave to enter the target's plot with a certain effect. This process of emplotment has turned out to be easier when applied to deceptions that have already occurred than to ongoing, fluid situations. It is also more a test of clinical acumen, than a didactic legacy.

In any case, the contextual approach will always have a place--even when a formal system has significant empirical support. This is because the formal approach excels mainly with bounded problems of inference that harbor known contingencies and discrete, dichotomous behavioral criteria (Levy, 1963). On the other hand, the contextual approach excels mainly with unbounded problems that contain unspecified or unknowable contingencies and continuous behavioral criteria. In fact, as far as each military deception is unique, the contextual approach may be the only way to predict whether a deception plan will work.

3. Military Deception Task Analysis. Even with a data base, psychological factors will affect how well deception planners do their job. Will they deceive or be deceived in the attempt? What ability, personality, and motivation characteristics are relevant to this question?

A pure task analysis which will generate behavioral criteria to which psychological information is associated cannot be performed. Unfortunately, implementation of a planned deception may be a low frequency event, confirmation of success or failure may often be lacking or far in the future, and planning behavior may be a product of bureaucratic politics more than past deception success. Thus, evaluating the predictive or postdictive validity of psychological information is problematic.

This leaves verbal effectiveness reports from peers and supervisors and data from nonmilitary, experimental deception tasks to which psychological information can be associated. Yet verbal reports are subject to political, personal, and knowledge biases, while experimental tasks have problems of ecological validity.

As an intellectual compromise, the following procedure has been used. A list of "important" deception tasks was obtained from the verbal reports of military deception planners. Experimental deception studies were abstracted which dealt with similar tasks. Personality characteristics associated with success in these tasks were abstracted from these studies. (Only studies with subjects having demograhic characteristics similar to military deception planners were used). Military deception planners were rated on these characteristics through the use of standardized tests with norms previously derived from samples of deception planners. Test scores were then correlated with task ratings per subject by peers and supervisors.

This admittedly crude and flawed procedure yielded four personality constructs positively associated with successful deception planning behavior. These were field independence as measured by the Group Embedded Figures Test (Witkin et al, 1971), integrative complexity as measured by the Individual Topical Inventory-Form A (Tuckman, 1966), Machiavellianism as measured by the Mach V (Christie, 1970), and empathy as measured by the Hogan Empathy Scale (Hogan, 1969).

In any official selection procedure, I would also include quick measures of intelligence, global psychopathology, and interpersonal style, as well as an interview and a relevant job history.

To my knowledge, no work has been done linking ability and motivation characteristics to military deception planning.

4. Foreign Deception in Psychological Research. Besides, generic factors, a data base, and profiles of successful planners, clinical psychology can be valuable in assessing the psychological research and development efforts of other nation-states. This assessment can include the issues of reliability, validity, and deception potential.

For example, there have been open-source reports of advances in voice stress analysis (One word's enough, 1982) and in "mind control" (Bloom, in press). The clinical psychologist's tolerance for ambiguity, analytic skills, knowledge in general psychology, and continuous awareness of conscious and unconscious deception can be virtues in avoiding the blinders of the true believer and the chronic skeptic.

For voice analysis, the lack of reliability and validity for the Psychological Stress Evaluator, the complexity of psychophysiological stress indices, subject control over acoustic parameters of the voice, and the dearth of replicable analyses of anatomical and physiological vocal parameters render claims of an operational and independent voice stress evaluator suspect (Horvath, 1978; Scherer, 1981).

For mind control, the clinical psychologist may hesitate less before suspicions arise--despite the often extraordinary psychotherapeutic feats of the likes of Milton Erickson. An analysis of similar concepts such as reflexology, spreading fields of activation, subliminal perception, and neurolinguistic programming renders an operational capability for mind control an opiate for the mindless (e.g., Bloom, 1983).

Conclusions. Let's close with a look at the status of each of the four clinical psychological contributions to military deception.

The five generic factors do not measure up to classical standards of reliability, but have been helpful as format aids to deception planners. The deception data base is still in the pilot study stages, but possesses tremendous import for the upgrading of military deception capabilities. The modified task analysis has yielded a selection procedure that is as close to operational as may be expected from the unique characteristics of military

deception planning. Evaluating the deception potential of behaviorally oriented research and development is an operational capability.

Hopefully, this brief paper will whet the appetites of others to delve into a complex, frustrating area--one that is intellectually stimulating and beneficial to national security.

REFERENCES

(Due to space limitations, references have been omitted from this edition of the paper. References are available upon request from the author at the address given at the beginning of the paper.)

DISCLAIMERS AND ACKNOWLEDGEMENTS

This paper does not necessarily reflect the views of the Armed Forces Medical Intelligence Center or Department of Defense.

The author is indebted to Mrs. Jeanne E. Hall for furnishing her outstanding administrative expertise.

AD-P003 240

Thought Control
From Parenting, to Behavior Mod, to Cults, to "Brainwashing"

Edna J. Hunter
United States International University

Ronald E. Hestand
St. Joseph State Hospital

Abstract

As long ago as 1957 Albert Biderman described the measures methodically used by the Chinese Communists in North Korea to induce compliance and "confessor" behavior in prisoners held by them. Biderman outlined a complex teaching procedure, gave illustrations of specific steps in the process and delineated the effects these measures had on the victim. Analyzing the procedures, it becomes evident that similarities run through a variety of situations where one faction makes a deliberate, intensive attempt to shape the attitudes and behaviors of another person under his or her absolute control. Whether it be the "all-powerful" parent, the abusive spouse, the intimidating military drill instructor, the Hell-fire-damnation religious fundamentalist, the terroristic kidnapper, or the POW interrogator, the techniques for inducing attitudinal and behavioral change ("brainwashing") and the victim's attempts to cope with these variations in stressors appear to be an analogous process.

Biderman's Eight Methods for Inducing Compliance

In a paper published by Biderman in 1957 on Chinese Communists' attempts to elicit false confessions from U.S. Air Force prisoners of war during the Korean conflict, eight coercive methods for eliciting individual compliance were enumerated. Biderman noted that the methods used by these captors were also commonly practiced by our own police and intelligence interrogators.

Most people equate "brainwashing" with actual severe physical torture, but such is not the case. For, as Biderman noted, "inflicting physical pain is not a particularly effective method." In fact, treating captives kindly, after first threatening death or physical harm, can be much more effective in eliciting subsequent compliant or cooperative behavior.

The eight techniques in the thought control process mentioned by Biderman included: (1) isolation of the captive; (2) monopolization of perception; (3) induced debilitation or exhaustion; (4) threats; (5) occasional indulgences; (6) demonstrating "omnipotence" or "omniscience" of the captor; (7) degradation; and (8) enforcing trivial demands.

The present report will examine this process in terms of how it is consciously or unconsciously applied to induce compliance in a variety of situa-

tions in our own culture today. For example, a similar process is employed in parenting, certain therapy groups, religious cults, and military training, as well as in more violent types of terroristic encounters.

Coercive Methods Within Families

Even the family is a source of coercive tactics, in many instances, which take place during the socialization period of the child and beyond. For example, almost from birth, the child is conditioned to respond quickly to the word "No." Not only do children learn "No," they are also quickly conditioned to believe they "can't do" this or that, regardless of whether or not the task in question is within their abilities. In most families, children are isolated from others until they reach school age, and even then, their associates are chosen carefully; there is monopolization of perception.

Children are often repeatedly told they are "bad" until they come to believe it. They are sometimes continuously bombarded with threats: "If you don't clean up your room, you're going to get it when your Dad gets home!" Although children have extremely short attention spans, they may be expected to perform far beyond their maturational abilities; thus, parents induce debilitation and exhaustion. For some children, there are long lists of chores that must be done, long periods of time spent in homework, or hours of practice at the piano, under threat of dire consequences if task performance does not meet parental standards. Certainly, from the child's viewpoint parents seem always to know what children are up to. Comedian Bill Cosby's monologues illustrate children's belief that parents have "eyes in back of their heads." It is not difficult to understand why children perceive parents as omnipotent and omniscient, just as the terroristic captor is viewed as all powerful by his captive.

How often parents scold children with words which imply the children are incompetent or unable to do even the simplest task without some form of aid. This type degredation is not far removed from that which Chinese Communists used in Korea to induce compliance.

The abusive family deserves special mention. Whether the abuse is physical, emotional, sexual abuse, or neglect, the same type techniques used in normal families are there, but in an extreme form. Abusing parents make trivial demands, expecting and promoting submissive behavior, are degrading, and have been shown to alternate between lenient and overly strict enforcement. It indeed sounds like "brainwashing."

Without elaboration, it bears mentioning that educational institutions, mental health activities, religious retreats, and even some therapy groups (e.g., EST), use the techniques of thought control.

Coercive Techniques and Religious Cults

Perhaps one of the most insidious forms of coercion is that used by religious cults, since cults are actually protected under the First Amendment of the United States Constitution.

According to Singer, the spread of cults began over 20 years ago, but the power of cult leaders was not recognized until the Guyana tragedy. It is estimated there are more than 2500 cults around that purport to offer mental, emotional, spiritual, psychological, and other benefits to their members (Singer, 1978). Usually cults are established by a strong or charismatic leader who controls followers by various means, all similar to Biderman's thought control techniques. Like the political prisoner or the prisoner of war, successful indoctrination of a cult recruit is likely to follow the same process, including isolation, control over communication, debilitation, degradation, induction of fear and confusion, alternation between harsh and lenient treatment, peer pressure, insistance that the individual's survival depends on belonging to the group, assignment of monotonous or repetitive tasks, and acts of symbolic betrayal of former values. The manipulation of cult converts is similar to the coercive persuasion techniques employed during the Korean War to achieve thought reform. Even among more traditional religious groups, one often sees religious revivals or retreats where the process of thought control techniques is clearly delineated (Kaslow & Sussman, 1982).

Coercion and Military Training

During Basic Training in the military raw recruits must be changed from individuals to "members of the team" -- members who do not ask questions, do not hesitate before obeying a command, but act immediately.

Little doubt exists that this form of training is effective, but by the same token, little doubt can exist that it is a form of coercive persuasion and follows the process Biderman delineated. From the moment servicepersons are sworn into the military, they are degraded or "put down," and made to feel inferior. The recruit is cloistered away with other recruits, and communication and contact cut off (isolation and monopolization of perception). Recruits are humiliated by having heads shaved, given conforming and ill-fitting clothing, and herded around like "animals." Recruits often report feeling "something less than human." They have lost control over their own lives and decision-making. They are likely to feel hopeless and helpless, that there is "no way out" -- much like the situations reported by prisoners of war and hostages.

The daily schedule in basic training is strenuous, leaving recruits utterly exhausted. Perception is monopolized to the point there is little time to think of pleasurable things. Threats, both veiled and blatant, are hurled at individuals who are already bombarded with feelings of being only "maggots" or "scum." Trivial demands are made of the recruit, demands for shoes that shine enough for drill instructors to see their faces in, for beds made so tightly a coin can be bounced off the covers, for ultra-close shaves, and for all types of insignificant requirements.

Drill instructors become demi-gods, omniscient and omnipotent, analogous to the POWs' perception of the captor or the abused child's view of the parent. Even officer training includes similar coercive techniques, as depicted in the recent motion picture, "An Officer and a Gentleman." Look at life at the military academies. During Plebe year, students are submitted to some of the same conditions as their enlisted counterparts during recruit training.

Students are "torn down and rebuilt" to the mold of what the military conceives as the perfect officer.

Terrorism and Compliance

The use of coercive persuasion again came to light with such notables as Patty Hearst, the Iranian hostages, and Leslie van Houten of the Charles Manson "family." Terrorism has been defined "to kill one person and frighten 10,000." It is a campaign of violence against "the system," waged from outside the system. Thus the rules of the system no longer apply. Civilians, of course, are not exempt as targets. One may be a target merely on the basis of nationality, ethnicity, religion, or pure happenstance, such as occurs when a bomb is placed inside a movie theater or in an airport.

Because terrorism appears indescriminate, it is more frightening and more difficult to protect against. But terrorism is not mindless, senseless, or irrational violence. It is a well- planned campaign of violence to advertise a group, to publicize its cause, and to create an atmosphere of fear and alarm.

It is very effective. The fear engendered by terroristic acts and the dramatic nature of them actually exaggerates the strength of the group. Terrorism is psychological warfare. The hostage is merely a tool to gain bargaining power.

We are somewhat accustomed to hearing about attacks on our embassies and on U.S. officials abroad; we are less accustomed to terroristic attacks within our own country. This unexpectedness is undoubtedly what accounted for the shock and fascination which occurred when Patricia Hearst was kidnapped by the SLA in 1974. Patty was carried, kicking, screaming, pleading "...not me, please, not me," the typical "this-can't-be-happening-to-me" response of most POWs, hostage victims, or other victims of traumatic events. After the kidnapping, for Patty there remained no semblance of security, only terror and uncertainty. We have only to examine Biderman's 1957 chart on the techniques of inducing compliance to see that Patty Hearst was in Phase I of the "brainwashing" process. A very real threat to life faced her. The second phase of the process is to induce distortion of perception through both physiological and psychological assaults. Patty was locked under a staircase, alone. There was no light, no other person, no way even to relieve her bodily functions, no sleep -- only fear, pain, and discomfort.

Once victims are "softened up" the manipulation begins. Physical terror is lifted and the process of de-identification with the captive's former environs commences, and re-identification with the captor emerges, just as it does during military recruit training.

Returning to the Hearst incident, finding onself in such a predicament, it is natural to attempt to establish blame. For Patty, first she most likely placed blame on herself; guilt would then ensue, just as it did when POWs were told they were "war criminals," and vulnerability to coercive persuasion increases. To rid herself of guilt, the next step was for Patty to project guilt away from herself, first to her parents and family, later to her entire social class. The re-identification had occurred.

When one becomes captive or hostage, the sentence is indefinite; the outcome unknown. Thus it is difficult to adjust to something that is completely ambiguous. Nonetheless to some extent the hostage or captive does adjust. Captives must cooperate with the captor if survival is to be achieved, at least to some extent. Typically, over time hostages, especially if held alone, will instinctively try to establish some "human bond" with the captor so they won't be killed. In the process, the captor sometimes may in turn become sympathetic with the captive. Dependent on the captor for the most basic of human needs makes for an almost parent-childlike relationship if the captivity lasts long enough. This countertransference was seen in the American hostages held in Iran, just as it has been in other captive/hostage situations in the past. Ironically upon release sometimes former captives will show more anger toward those who rescued them than toward the captor who held them (Hunter, 1983).

Conclusions

In this paper we have pointed out how coercive psychological techniques are prevalent in our society, not only in malignant situations such as POW and hostage experiences and terroristic kidnappings, but also in relatively benign situations such as religious retreats and childrearing practices. Analogous methods are also applied in advertising, in our educational systems, in therapeutic situations and behavioral modification efforts. It is important to reexamine these techniques emphasizing the ethics and human rights involved, particularly in those areas where we have some control. Certainly, within both civilian and military life, individuals differ in their susceptability to thought control. Perhaps if all of us were not continually conditioned from birth onward to bend to others' value systems in almost every institution we enter, coercive persuasion might be easier to resist. However, we must remember that coercive psychological techniques, like research findings, can be used for good as well as malevolent outcomes. Sometimes we walk a thin ethical line between the two.

Nonetheless, there is something we can do for the children and military members of tomorrow. Research has shown us that individuals are far less susceptible to coercive persuasion where there is group support from comrades, where there exists a well-ingrained sense of values, and where the individual has a strong commitment to a cause, be it self, God, or country. Thus, we are back to a recognition of the critical importance of early childhood training and skillful, effective parenting.

References

Biderman, A.D. Communist attempts to elicit false confessions from Air Force prisoners of war. Bull. N.Y. Acad. Med., Sept. 1957, 33 (9), 616-625.

Hunter, E.J. Coercive persuasion: The myth of free will? Air University Rev. Jan-Feb 1983, XXXIV (2) 100-107.

Kaslow, F. and Sussman, M. (Eds.). Cults and the Family. New York: The Haworth Press, 1982.

Singer, H.T. Coming out of the cults. Psychology Today, 1978, 12 (8), 72-82.

Relevant and Irrelevant Legal Structures: Distinguishing Private Sector from DOD Contracting

Edward J. McCabe
Department of Sociology
Oakland University
Rochester, Michigan

Abstract

Based on interview and observational data of contract formation and administration dealing with hardware and publications development by the U.S. Army Tank-Automotive Command, the social process of government contract law as an example of law in action is compared with the typical use of contracts between merchants in the private sector. While the formal law of contracts is, for the most part, irrelevant in normal business exchanges, the formal legal structure is found to be routinely relied upon in the case of government contracts. This unusual role of the formal law is explained by the absence of normal reciprocal relationships between contractors and the Government.

Introduction

Behavioral scientists have long been interested in reciprocity and exchange relationships. One of the societal mechanisms for ordering relationships is the law and an enormous amount of the exchange in American society falls within the broad scope of contract law. In addition, many of those exchanges involve governmental entities as contracting parties.

During the last fiscal year, governmental entities in the United States purchased over one-fourth of the gross national product. All departments of the Federal Government and most of the major agencies conduct significant procurement programs in which Government contracts with suppliers of goods and services in the private sector. Historically, the Department of Defense and the General Services Administration have been the leading agencies in promulgating broad procurement policy with DOD accounting for the largest volume of acquisitions. The formal body of law consisting of the relevant statutes, regulations, orders, and judicial decisions is known variously as "procurement law," "public contract law," or "the law of government contracts."

In American law, contracts are legally enforceable promises through which parties may privately order their relationships and by which parties may invoke the courts to either force compliance or recover damages for breach. The power of the United States to contract is incident to the general powers granted by the Constitution and upon entering a contract the Government becomes subject to the rule of Federal law as a private individual.

A legal system may often be discussed as if it operates entirely in harmony with the way in which it is described or prescribed in constitutions or legislative acts, glossing over departures from idealized conceptions. Many tend to assume a naive unity between the formal model of law and the way it works in reality. However, in order to understand legal relationships and the nature of rights and duties as components of behavioral reality rather than mere prescriptive ideals, it is not sufficient to only know the formal rules. One must understand the law in action. "Inquiry [should be] on what human beings, cast in socially defined roles in certain characteristic types of decision-making sequences which have traditionally been identified as 'legal' do in their interaction and transaction with each other." (Schubert, 1968, p. 408).

The discussion below attempts to broadly sketch the usual pattern of interaction of those nominally engaged in contractual relationships. Comparisons are made between the general patterns of relations found in the private sector "between merchants" and those observed and described in a limited area of government contracting.

Method

Comparisons are made on the basis of current, social science understanding of contract-based relations in the private sector "between merchants," (see, for example, Macaulay, 1963), and observation and interview data describing contracting and contract administration relating to limited types of DOD contracts. Contracts "between merchants" are those private sector transactions in which the parties are chargeable with the knowledge or skill of those who normally deal in such goods, i.e., the "normal commercial practices" between private businesses. The relevant government contracts were limited to a selection of those dealing with hardware and publications development, acquisation, and maintanence processed through the Department of the Army, U.S. Army Tank-Automotive Material Readiness Command. Individual interviews and group discussions were held over a 14 month period with a total of 87 TACOM employees representing individuals with a variety of contracting and contract administration experience.

While it is difficult to generalize about the law in action as it applies to TACOM contracting, it is believed that a number of observations can be made with reasonable accuracy at this time. Nonetheless, it important to note that all conclusionary observations are the sole responsibility of the author.

Findings

Contracts are devices used for conducting exchanges. Such devices are not synonymous with the exchange itself which may or may not be characterized as contractual. In this sense, to contract involves both rational planning of the transaction with careful provision for as many future contingencies as can be foreseen and the existence of actual or potential legal sanctions to induce performance of the exchange or to

compensate for non-performance.

These devices for conducting exchanges may be used or may exist in greater or lessor degree. As a result, transactions can be described, relative to each other, as involving a more or less contractual manner (a) of creating an exchange relationship or (b) of solving problems arising during the cause of such a relationship.

The Creation of Exchange Relationships. Parties negotiating a contract can make plans concerning several types of issues: (1) They can define what each party is to do or refrain from doing. (2) They can plan what effect certain contingencies are to have on their duties. (3) They can plan what is to happen if either of them fails to perform. (4) They can plan their agreement so that it is more explicitely legally enforceable. On each of these issues, the level of planning that the parties actually engage in may range from careful and explicit planning of an issue to simply being unaware of the issue.

While many exchanges between merchants reflect a high degree of planning about each of the four categories of issues, many, if not most, exchanges reflect no planning, or only a minimal amount of it, especially concerning legal sanctions and the effect of defective performances. (Macaulay, 1963). Typically, in the private sector, businesses pay the most attention to describing the performances in an exchange and this level of planning may be little more than sending an order for a specified quanity of goods and relying on "trade usage," "past dealings," or assumptions like "Smith will know what we need" to supply missing specifications. There is usually less planning for contingencies, even less for the possibility of defective performances, and, typically, little explicit recognition is given to making their contracts legally enforceable or to providing specific legal sanctions.

Article 2 of the Uniform Commercial Code (American Law Institute, 1978), which has been adopted in every state except Louisiana, recognizes such typical practices. Under Article 2 of the U.C.C., other things being equal, a contract will be valid and enforceable between merchants so long as a reasonably specific quantity term is included in the agreement. The Code will supply other missing terms through various "gap-filler" provisions. Although there are over 4000 statutes and executive orders that are relevant to the Federal Government contract process (Department of the Army, 1983), the Uniform Commercial Code and its gap-filler provisions does not apply except as dicta, i.e., non-binding judicial opinion. Missing terms will not be supplied except through an addition or modification to the contract.

In sharp contrast to typical contracts between merchants, exchanges based on DOD contracts usually represent a high level of planning in each of the four areas. The typical government contract seems to "have a clause for everything and everything in its clause." There are explicit statements concerning each of the four categories: performance, contingencies, defective performances and legal sanctions. On the other hand, compared to some ideal standard of rationality, much of this planning is illusionary, at least as far as each specific contract is concerned. While specific

performance is usually the result of a high level of individual planning, most other issues are dealt with through the selective addition items from a menu of "boilerplate" provisions. Nonetheless, relative to typical business exchanges where such issues may be virtually ignored or appear only in the exchange of mutually contradictory order and invoice froms, government contract "boilerplate" represents a comparatively high level of planning.

Dispute Settlement and the Adjustment of Exchange Relationships. While a large amount of creating exchanges between merchants is done on a relatively non-contractual basis, the creation of business exchanges is usually far more contractual than the adjustment of such relationships and the settlement of disputes. As noted above, the creation of government contracts is far more contractual than the creation business exchanges. While modification of relationships and the settlement of disputes concerning government contracts is more contractual than the adjustment of business exchanges, these are less contractual than the initial creation of such relationships.

In the private sector, business exchanges are usually adjusted informally. Even where the parties have a detailed and carefully planned agreement that has specified what is to happen in the event of the now relevant contingency, often parties will never refer to the agreement but will negotiate a solution when the problem arises as if there had never been an original contract. Many believe as the businessman quoted by Macaulay (p. 62), "You can settle any dispute if you keep the lawyers and accountants out of it." There are some important exceptions to generalization noted below.

In government contracts, necessary modifications or disputes are much more likely to be handled in a formal legalistic manner specified in the applicable contract provisions or relevant statute. Nonetheless, consistent with expectations, the law is action is often more pragmatic and subject to human intervention, consideration, and negotiation than would first appear upon inspection of the formal rules. For example, a frequent estimate by respondents was that 60 to 75 percent of the contracts with which they were familiar in recent years where technically in default at one time or another. A contract can be technically in default, and as a result the contractor can be subject to a show-cause order prior to the withholding of any scheduled payment, if any product, report, or form is 30 days late. However, very few such orders are ever issued.

Discussion

Although the general public appears to overestimate the relevance of contract law in normal business relations, it is not surprising that relatively non-contractual practices are so common in the private sector. In most cases the possible functions of contracts are served by other devices. Most problems are avoided without resort to detailed planning or legal sanctions because the parties understand the primary obligations and they are involved in an on-going reciprocal relationship which they wish to maintain. The maintanence of such a relationship provides strong incentive to settle disputes, even at perhaps a short-term disadvantage. For example, during the recent depressed sales in the auto industry, General Motors

unilaterally announced to its suppliers that the company would pay 10% less for all of those supplies for which it had previously contracted. The suppliers may have privately grumbled but they did not file suit.

Within the private sector, there are two main exceptions to the normal irrelevance of formal contracts. Both involve circumstances in which there is a lack of a continuing relationship. The first situation is when parties enter into a single, transient exchange with little likelihood of repeat business or a continuing relationship. Sales of small businesses from one owner to another, home and auto sales, home remodeling contracts are illustrations. The second situation is when a previously continuing relationship is terminated. An example is an action based on the termination of a dealer's franchise by a manufacturer. Since the franchise has been terminated, the dealer may have little to lose in sueing the manufacturer, especially under a contingency fee arrangement.

At first, it would appear that the situation of the Government and its defense contractors is dissimilar. Especially with major weapon systems, the Government repeatedly deals with the same relatively few prime contractors. However, unlike private parties, the Government is generally prohibited from using the knowledge gained in such a relationship in future considerations. While it is possible for a contractor to become disqualified from receiving further awards based on poor performance under previous contracts, this is extremely unlikely except in flagrant circumstances. If a previously marginal performer is low bidder, all things being equal, the Government may not reject the contractor. Further, many of the techniques by which those in continuing relationships may influence each others behavior, such as gratuities, future considerations, or even forms of extortion, are prohibited. When commentators advocate bringing business efficiency into government, they may also be inadvertantly advocating that government engage in illegal activity.

REFERENCES

American Law Institute. (1978). Uniform commercial code (9th ed.). Philadelphia: Author.

Department of the Army. (1983). Legal services procurement law. Washington, D.C.: Author.

Macaulay, S. (1963). Non-contractual relations in business. American Sociological Association, 28(1), 55-67.

Schubert, G. (1968). Behavioral jurisprudence. Law and Society Review, 2, 408-418.

NOTES ON A THEORY OF TERRORISM

Maynard Shelly &
Felix Moos
University of Kansas

Terrorism is expected to be a growing problem in much of the world. But as yet there is no theory that covers the phenomenon as the complex and dispersed multi-faceted activity that it is. We are developing a theory of terrorism we hope will encompass much of its complexity. This paper summarizes some of the developments in this theory. We make no attempt to include references since these are extensive and referred to in some detail in other papers and manuscripts. In this paper we emphasize the motivation for going in the directions we have taken.

A theory of terrorism is of interest to the military for two reasons. The first is that the U.S. Armed Forces, both at home and overseas, are evermore frequently targets of terrorists. In some ways, the American military symbolizes the strength of America and may therefore be attacked to demonstrate that the United States is weaker and more vulnerable than it appears to be. The second is that the military, in the form of the National Guard, may be called upon to secure a domestic area when disruption exceeds tolerable levels. Both authors have call to remember this, since both of us helped guard buildings at the University of Kansas in the 1960's, when the Kansas National Guard was called out in Lawrence, Kansas following the burning of the University's Student Union. Further, we assume that because of the totally destructive effects of a nuclear war, terrorism has become a "least-worst military option". Thus, terrorism has in fact developed into an acceptable surrogate warfare.

Two Aspects to Any Theory of Terrorism

Any theory of terrorism must consider at least two aspects. We call these the micro- and macro-aspects of terrorism, terms taken from economics. Micro-economics refers to models describing the behavior of the individual consumer or firm. Macro-economics refers to models of an entire economy. The micro-aspects of our theory cover the dynamics of individual terrorists. Its macro-aspects describe the encompassing social circumstances in which terrorism develops. A complete theory of terrorism, which we are attempting to develop, will combine the two aspects into a single model.

Although we feel that the problems of terrorism are very complex and that any model describing it must respond to these complexities, most approaches to the problem have not been very sophisticated. Generally existing approaches fall into four categories. First is the "gossip collumnist" approach exemplified by Claire Sterling. This approach is useful to those interested in specific

groups, but it does not describe conditions conducive to terrorism; does not identify the dynamics of individual terrorists, and even less provides strategies for dealing with it.

The second approach is more useful. It attempts to describe the motivations of individual terrorists. A book like F.S. Hacker's Crusaders, Criminals, Crazies can be useful. This type of book and related papers permit those who must deal with terrorists to achieve a somewhat better understanding of them. Generally, ho-.ever, these approaches exclude some of terrorist motivations we consider to be important, and fail to consider the interactions of personal characteristics with societal circumstances (national characters). Nor do they consider the effect of particular cultures in structuring or modifying a terrorist's motivations.

The third approach focuses on particular sites or episodes and is most useful to those who must react to terrorist actions. This approach describes the hardening of targets, hostage negotiations, and reactive patterns that are most effective in minimizing a terrorist threat. The usefulness of this approach, however, is reduced by the relative unpredictability of terrorist targets and the intervals between attacks. As has been pointed out repeatedly, when one target is sufficiently hardened, terrorists may simply revert to another less hardened target.

The fourth approach is the most useful, It examines strategic problems instead of the tactical ones focused upon by the third approach. The terrorist's strategies and methods are examined with a view to what might be used to counter their strategies. This approach points to the importance for good intelligence. Both this and the preceding path to understanding terrorism leave behind whatever contributions the second approach might make, that of comprehending the dynamics of individual terrorists. Although this last approach is probably the most important, it is also the one that is least well developed. Issues are dealt with largely eccletically rather than in an integrated way. For example, the problems of good intelligence and the proper reaction to a terrorist organization trying to force an overreaction by authorities has not been covered as far as we know.

The Micro-Theory

We cannot describe the micro-theory that is being developed in great detail. We can present, however, some of our motivations and reasons for our approach. In particular we are seeking a simple structure than can depict a wide variety of behaviors and be conveniently described by a system of linear differential equations. This latter attribute is necessary to integrate our micro- with our macro-theory of terrorism.

In the micro-theory we describe what we call for convenience, six "forces" motivating individual terrorists. In selecting these six characteristics we have used one overriding criteria. Every "force" must have a basis in at least biology, and preferably throughout the harder sciences. Such a criteria will ensure that the aspects we are dealing with are not only convenient but probably fundamental as well.

Basic to our micro-theory are the attractive and repulsive forces. Attraction and repulsion are found throughout nature from the behavior of inanimate bodies

to human beings; the ability to be attracted and repelled form the basis of modern behavioristic theory. The basic quality of the attractive force is seen in people who find nothing attractive; they are so depressed that they can accomplish nothing.

Of next greatest importance are the growth and conservative forces. All complex living organisms seek new experiences. For some predatory animals moving from terroritory to terroritory (curiosity) has the advantage of not killing off the game in a region. We call this the growth force because it expands the animal's set of memories and prepares it to deal with a wider variety of circumstances. By contrast the conservative force leads an animal to seek the familiar and withdraw from the excessively strange. The conservative force is a way of grouping activities which tend to preserve the animal's life.

At the top of the hierarchy are the most fundamental, as opposed to the most basic, forces, the creative and destructive forces. While the attractive and repulsive are basic to all the other forces in that the other forces depend upon them (such as in being attracted to the familiar), the creative and destructive forces reflect what we believe to be the most fundamental aspects of nature--its tendencies to move towards increased meaningful orgnization and destruction or disorganization.

There are many examples of the tendency of nature to move towards greater organized complexity in the universe including the development of solar systems, galexies and life itself as well as the more contemporary examples of increasing technology and the growing sophistication of communication systems. Disciplines and approaches have arisen to try to understand this phenomenon better including synergetics, dissipative structure theory, and the descriptions of nature based on Jantsch's "evolutionary vision". Sometimes this tendency towards greater meaningful organization is called negentropy.

There are equally many examples of the tendency of nature to move towards less orgnization as in the collapse of stars, the dissipation of heat, the decay of dead bodies, and some aspects of economic processes. These processes have been called entropic processes and have been better described than negentropic processes.

In the human being we theorize that the most fundamental force of all is that of the creative force which manifests itself in the desire to learn, in an insistent need to know why someone dislikes a person, and in a general tendency for people to become increasingly cognitively complex as well as in the perpetual quest for religious understanding. Almost as important is the destructive force which spawns an interest in everything from horror movies to combat and stimulates attacks against others as well as suicide. No understanding of human behavior and experience, including in particular that of terrorism, is possible without relying upon these forces. It is our reliance upon these aspects of human beings that sets our micro-theory apart from that of many other extant psychological theories.

We assume further that the culture in which an individual is socialized most often compels him/her to make continuously culturally congruent choices. To know something of the specific culture of the terrorist is to know something about which possibilities a person will recognize and which one will be most likely selected for action. To know the various themes, traits, configurations

and specific patterns of a culture, and to understand the nature of the dynamics encompassed by a particular "national character" produced by the interplay between individual (forces) and culture, is to know a great deal about the possibilities that will be recognized and utilized by the individual terrorist and/or terrorist organization. Although all terrorists engage in destructive acts, the choice of act, the location of the act, and the method by which the act is carried out will tend to be different for the Palestinian, Irish, German, or Japanese terrorist.

In our approach there is no single type of terrorist. Instead, different terrorists are characterized by the forces that are strong in them. They, as well are characterized by their differing cultural antecedants and national characters. For example, evidence suggests that the infamous terrorist "Carlos" is dominated by the growth force while the assassins of Anwar Sadat were dominated by the conservative force (and hence by a desire to destory that which threatened to bring too much change). Just as Carlos behaved as a "Latin" in his activities so, to a great extent, did the murders of Sadat, behave as fundamentalist Arabs.

(We are using Monte Carlo simulations to see how these six forces might be interrelated.)

The Macro-Theory

At present the macro-part of our theory of terrorism is less well developed. Basically, we need a model which is capable of describing activities ranging from the acts of terrorists to the reactions of citizenry to terrorist episodes. Such a model could come from anthropology, sociology, political science or economics to provide our basic model since it is more developed, and because it is already divided into micro- and macro-theories.

We begin with a model of the abstract economy and enlarge it to include aspects essential for theorizing about terrorism. While the length of this paper does not permit description of the model itself, we can present some of its basic assumptions in overly simplified form.

The most important assumption that we make in the macro-theory is that an expanding society (such as one with an expanding economy) tends to have little or less terrorism. This is because we believe that there is an "urge" in societies as well as individuals to progress towards greater organized complexity. When this urge is satisfied, as in Europe and the U.S. in the 1950's and most of the 1960's, there tends to be little terrorism.

Another assumption, closely related to the preceding, is that an economic or political system that does not become increasingly complex deteriorates. This was seen in Spain during the Franco regime and in China under the Red Guards. Efforts to maintain such a state will result in an apathetic society or in terrorism or rebellion. Terrorism may be most likely when one part of the population feels themselves to be inequitably treated. Examples include El Salvador, Northern Ireland and Lebanon.

Finally, we assume that (from the theory of dissipative structures and the applications of entropy theory to economics) progress creates entropy in

addition to creating more complex meaningful social structures. Part of this entropy or disorganization becomes expressed as terrorism. The best example that we can think of to illustrate this idea is Iran. There the fundamentalist followers of Khomeini objected to the modernization taking place because it was destroying old customs without providing enough time to acquire a new culture.

The Complete Theory of Terrorism

The complete theory is far from finished. For example, we have to find ways of including as part of the interaction between the two theories the influence of culture and national character. Here we give only two examples of the interaction of the two aspects of the theory, one in which both theories reinforce the same conclusion and another example in which the micro-theory modulates a assumption of the macro-theory.

The axiom that destructiveness increases whenever any of the forces are thwarted is part of the micro-theory. Since the creative force is the most fundamental of the forces, its being thwarted will produce destructiveness including terrorism. A society growing in meaningful complexity will offer individuals opportunities to become more complex and satisfy the urgings of the creative force. But when the progress of society slows down, these opportunities diminish. An example is the fact that most of the originial Red Brigade members in Italy were Humanities graduates who found insufficient opportunity to take advantage of their college educations in an increasinly technological world. This is why, we believe, universities are often origins of political violence; the creative force is excited but (especially in poor nations) not given sufficient opportunities for its continued developments.

We can refer to the late 1960's in the U.S. as an example of the micro-theory modulating the macro-theory. Although the U.S. economy was still expanding groups like the Weathermen (later the Weather Underground) developed. This should not have occurred according to our macro-theory (in its simplified form). We believe that the growing sexual freedom and stimulating effects of the emerging life-styles of that time created a growth force that could no longer be adequately satisfied. (When the stronger growth force leads the person to understand or seek meaning in his or her array of new experiences, the axiom about a growing economy producing little violence is better observed.) It is difficult to provide enough novelty for people who are constantly fighting boredom. In such cases, as with gangs in the center of many U.S. cities, violence is always a final option for avoiding boredom. In this way we see how the macro-theory is modified by the micro-theory.

Conclusion

We have described some of our motivations and a few features of the theory of terrorism that we are developing. As we continue to work in the area it becomes increasingly clear to us that theories providing a useful reaction to terrorism will have to be a great deal more complex than most of those being proposed at present.

Situational Interaction: A Peer Counseling Approach to AWOL Reduction

Robert L. Durham and Carol Emilio
University of Colorado, Colorado Springs
and Captain Paul Barkey (Chaplain)
Fort Carson, Colorado

Abstract

An AWOL reduction program, which utilized peer counseling, was conducted with two randomly selected companies from a mechanized infantry battalion at Fort Carson, Colorado. Two randomly selected companies from the same brigade served as static control units. Employing the Taylor-Johnson Temperament Analysis (T-JTA), AWOL-prone soldiers were identified and were counseled initially by the unit chaplain and subsequently by platoon leaders. Platoon leaders identified situational aspects of AWOL-prone soldiers and interacted as mediators between environmental situations (e.g., money problems) and personal factors identified by the T-JTA. As a function of the intervention, the treated group showed a significant decline in AWOL rates (t (24) = 3.07, $p < .01$) while the control group did not. Results were discussed in terms of the efficiency of employing the Taylor-Johnson Temperament Analysis in conjunction with peer counseling to reduce AWOL rates.

Introduction

Unauthorized absences from duty (AWOLS) have existed for as long as there have been organized military services. Past research (e.g., Bell & Holz, 1975; Drucker & Schwartz, 1973; Hartnagel, 1974) has demonstrated the existence of a relationship between AWOLS and volunteer members. That is, it appears that more volunteers than draftees go AWOL. However, the major consequences to the armed services are the same: reduction in combat capability, lowered morale of those who are present for duty, and the cost of documenting and processing the absentees.

Historically, research on military delinquency has focused on the development of predictive instruments based on personality, background, and demographic variables (i.e., individual variables). The goal was to identify individuals with potential disciplinary and delinquency problems either before they entered the service or while they were in basic training (Bell & Holz, 1975; Berbiglia, 1971; Littlepage & Rappaport, 1977). A predictive device would enable the service to reject recruits who would most likely have delinquency problems.

Measurement devices that have been utilized to identify potential AWOL personnel include the Background and Opinion Questionnaire-72 (BOQ-72), the Taylor-Johnson Temperament Analysis (T-JTA), and the Demographic Questionnaire (DQ). In 1973 and 1974, the BOQ-72 was used at basic combat training (BCT) units at Fort Polk, Louisiana and Fort

Knox, Kentucky as a means of identifying possible AWOLs. However, preventive programs based on the BOQ results were not successful. Bell and Holz (1975) have suggested that the use of commanders as interviewer-counselors may have contributed to the outcome. First, due to the constant turnover of commanders, the program lacked consistency. Lack of consistency was also suspected in the counseling process itself because, although the commanders received outlined instructions to follow in the counseling of troops, procedures were not utilized to determine the extent of application. Most importantly, however, individuals were aware of having been labeled as possible delinquents which may have led to the self-fulfilling prophecy effect.

Berbiglia (1971) utilized the T-JTA at Fort Polk, Louisiana. He had observed that soldiers confined to the stockade for AWOL offenses scored consistently differently on the T-JTA than soldiers confined for non-AWOl related offenses--a phenomenon he labeled "AWOL Syndrome." This information was subsequently used for identifying potential AWOLs _a priori_. Although statistics were not available, Berbiglia indicated that the AWOL rate could be decreased by 50% when identification of potential AWOL candidates was followed by counseling, and, if necessary, referral to other post agencies (e.g., mental hygiene, finance, legal services, Red Cross).

The ARI evaluated a later program at Fort Polk which used the T-JTA in conjunction with the BOQ-72 to determine potential AWOLs. All subjects selected for the study took one of the tests, but one-third of them were not scored and were used as control subjects. The individuals identified as possible delinquents were then subjected to a modified preventive program where the company commanders again provided the counseling and referral services (Bell & Holz, 1975). According to the ARI, the T-JTA did not reliably predict AWOLs and the process of identification actually increased the AWOL rate.

Recently, Bell and Holz (1975) and Littlepage and Rappaport (1977) have proposed an interactive model as an alternative to predictive devices. According to their perspective, emphasis should be placed on the environmental and situational factors (i.e., problem events and organizational characteristics) of the miiltary service itself, since these components interact with personal factors and can be changed relatively easily while individual factors cannot. Some of the organizational characteristics have been identified as problem solving help, leader consideration, job satisfaction, unit atmosphere, and leader ability.

Although an interactive model of AWOL behavior appears to be the best explanatory device at the present time, its use has been limited to either confinement installations (e.g., Littlepage & Rappaport, 1977) or basic training settings (e.g., Bell & Holz, 1975). It must be emphasized that the use of the earlier singular predictive models have also been limited to these same settings--a situation which may have external validity implications. As Drucker and Schwartz (1973) pointed out, the causes of AWOL in basic training probably differ from the causes of AWOL in regular assignment stations.

The present study was concerned with the evaluation of the effectiveness of an AWOL reduction program based on the interactive model. After reviewing Berbiglia's use of the T-JTA for identifying an AWOL syndrome, the battalion chaplain proposed to the battalion commander a program of testing, intervention, evaluation, and follow-up. The commander and chaplain then presented the program to the unit leaders with emphasis on the principle that most intervention procedures would be done by small unit leaders at the platoon level.

Method

Subjects. The experimental and control groups constituted a non-random sample of soldiers who were stationed at Fort Carson, Colorado. Two companies of the same disposition from a mechanized infantry battalion served as the experimental unit. Equivalency of the control subjects was achieved by using soldiers from identical companies in a different mechanized infantry unit in the same brigade. There were approximately 720 subjects total from the two battalions.

Procedure. Testing of the experimental units via the T-JTA was done one company at a time and by platoons, including platoon leadership over a one-month period (i.e., soldiers entering the battalion after this period were not tested). The test was administered by the chaplain and the instructions included a statement that no one would be allowed to see the results of the individual tests except the individual and the unit chaplain. However, some information would be used by the chaplain and unit leaders for administrative decisions. The actual purpose of the test was not revealed in order to help prevent contamination, self-fulfilling prophecy effects, and response bias. The chaplain screened all tests personally and initiated interviews with persons exhibiting extreme scores on four or more of the T-JTA scales describing an individual as nervous, depressive, hostile, quiet, inhibited, or impulsive (i.e., the AWOL Syndrome).

The chaplain then met with the platoon leaders and company commanders to discuss how to relate to individual soldiers based on test results and the problems the chaplain encountered during the initial interviews. This led to the development of a strategy, based on individual need, for dealing with personal problems by manipulation of organizational aspects of the immediate environment, at the unit level, where possible. Once the strategy had been developed, unit leaders became responsible for its implementation with the chaplain functioning mainly as a resource or back-up person. For example, if the T-JTA subscores revealed tendencies toward depression, nervousness, hostility, and impulsiveness (personal factors) and it was further determined during the chaplain's interview that this person was experiencing financial difficulties (a problem event), financial aid would be sought through such channels as the Emergency Relief Agency or Red Cross. A follow-up session would then be conducted to ensure that the problem had been taken care of and to determine if further assistance was needed in other areas.

AWOL data were then gathered for the 14-month period prior to the intervention and the 12-month period following it. The data were then

compared to a control group, for the same period of time, which received neither testing nor intervention.

Results

Utilizing ARIMA (Auto-Regressive-Integrated-Moving-Average) methodology for this purpose, the AWOL data were found to be independent. That is, the data fit the ARIMA (0,0,0) model. The lag one autocorrelations for the pre- and post-intervention experimental data were -.10 (X^2 (6) = 4.72, p > .05) and .22 (X^2 (5) = 3.44, p > .05). Pre- and post-lag one autocorrelations computed for the control group were -.05 (X^2 (6) = 1.16, p > .05) and .16 (X^2 (5) = 2.38, p > .05) respectively. These results demonstrate the independence of the data for both groups justifying the use of t-tests for independent data (Glass et al., 1975).

The mean number of AWOLs for the experimental and control groups did not differ significantly in the pre-intervention condition (t (26) = .23, p > .05). However, the experimental group demonstrated a significant decline in mean AWOL rates between the pre- and post-intervention conditions (t (24) = 3.07, p < .01). The control group did not demonstrate such a decline (t (24) = 1.34, p > .05). The mean number of AWOLs per month for pre- and post-intervention conditions for the experimental group were 5.21 and 2.58 respectively. The mean number of AWOLs per month for pre- and post-intervention for the control condition were 5.0 and 3.66 respectively.

Discussion

In order for an intervention program to be effective, the type of intervention must meet some minimal standards of relevance to the existing problem. This must be so not only in substance but also in approach. AWOL-prone individuals could interact not only with the chaplain, but also, more importantly, with persons who were closer to the AWOL as peers. Discussions could therefore have been viewed as more personal, relevant, meaningful, and helpful to the potential AWOL, especially within the context of manipulating (changing) situational factors. For example, the chaplain noted that on many occasions platoon leaders returned to the chaplain for advice on dealing with specific problems. Thus it can be seen that the present intervetion program utilized first, the identification of AWOL-prone soldiers through the T-JTA; second, it provided a means for counseling not only through the chaplain's office but also at the small unit level with peers; and third, it allowed the platoon leaders access to the chaplain for advice in counseling. These three components contributed to the effectiveness of the program.

Three final points need to be emphasized with respect to this project. First, it should be noted that, upon completion of the T-JTA testing at the onset of the program, subsequent soldiers transferring into the units were not tested. Thus, any new transfer individuals were not identified as potential AWOLs. The fact that the experimental units showed a decrease in AWOL rate in spite of this added threat to the efficiency of the program indicates the strength of the intervention treatment.

Second, it should be pointed out that the current project is one of the first to demonstrate an AWOL reduction program with statistical documentation. This is important, as it indicates the reliability of the T-JTA used in conjunction with peer counseling for the reduction of AWOLs in a regular unit. Again, it should be stressed that this particular treatment intervention is probably not generalizable to other types of units (i.e., basic training units), as the problems encountered by AWOL-prone individuals in such units are different from those at regular duty stations.

Third, it should be noted that the present investigation does not differentiate the obtained experimental effects with respect to the T-JTA and peer counseling. That is, the results could reflect the effects of counseling alone, the interaction of counseling and testing, or the effects of testing. The latter possibility is unlikely as the soldiers were blind to the reasons for testing.

References

Bell, D. B., & Holz, R. F. 1975. Summary of ARI research on military delinquency. (ARI Research Report 1185). Arlington, Va.: U.S. Army Research Institute for the Behavioral and Social Sciences, June.

Berbiglia, J. C. 1971. The AWOL syndrome: A study in the early identification of potential AWOLs by use of the Taylor-Johnson Temperament Analysis leading to the development of a preventive program. Psychological Publications, 3, 1-20.

Drucker, E. H., & Schwartz, S. 1973. The prediction of AWOL, military skills and leadership potential. (HumRRO Tech. Rep. 73-1). Alexandria, Va.: Human Resource Research Organization, January.

Glass, G. V., Willson, V. L., & Gottman, J. M. 1975. Design and analysis of time-series experiments. Boulder, Co.: Colorado Associated University Press.

Hartnagel, T. F. 1974. Absent without leave: A study of the military offender. Journal of Political and Military Sociology, 2, 204-220.

Littlepage, G. E., & Rappaport, L. 1977. Factors affecting military AWOL decisions. Journal of Political and Military Sociology, 5, 117-125.

Evaluation of the BCT Paraprofessional Counselor Training at the United States Air Force Academy

Paul R. Brown
United States Air Force Academy

Jarean L. Ray
USAF Occupational Measurement Center

Abstract

This study evaluated the competence of USAF Academy paraprofessional counselors in communicating a helpful response to assess the success of counselor training. Fifty-eight of the 71 trainees reached a communication competence level of 2.0 or higher, indicating they will not cause harm to the client in counseling.

Introduction

The purpose of this study was to evaluate, by objectively measuring counseling ability, training for paraprofessional counselors who worked as cadre in Basic Cadet Training (BCT) at the United States Air Force Academy. BCT is a six-week training program for cadets entering the Academy, designed to provide military instruction and experience to transition the basic cadet from civilian to military life and provide a foundation for future military development. To meet the individual-specific needs caused by the strenuous environment, the USAF Academy uses upperclass cadets as paraprofessional counselors to the basic cadets.

An important consideration in the use of counselors is his or her effectiveness in counseling. Egan (1982) emphasizes this consideration by stating that "unskilled and mismanaged helping can do a great deal of harm." However, with appropriate training, lay helpers can be brought to effective levels of counseling. Carkhuff (1969) states:

> A review of these lay training programs indicates that (1) lay persons can be trained to function at minimally facilitative levels of conditions related to constructive client change in relatively short periods of time and (2) lay counselors can effect significant constructive client change in clients.

Successful training involves bringing the trainee to a minimum level of counseling competency so he or she will, at the least, not cause harm to the client. Of course, the desired outcome is for the trainee to be facilitative.

Although there are many approaches to training counselors, the Academy uses a microcounseling (MC) approach. MC involves dividing the desired counseling skills into basic components that can be taught in a sequential process. Brammer (1979) mentions several advantages of MC, including a flexibility that allows adaptation to different training requirements and connection between theory and practice. Also, current research supports the effectiveness of an MC approach; for example, Ford (1979) cites many studies

that support the greater effectiveness of an MC approach over several other approaches. Anectodal data concerning the success of the paraprofessional program in helping basic cadets adjust to the Academy (Brown, 1981) suggests that this training program is effective; however, this study is the first attempt to evaluate the training program.

Carkhuff (1969) explains that both discrimination and communication of a facilitative response are necessary for the counseling to be effective. Discrimination involves the helper's ability to perceive that a certain response is constructive, rather than harmful. Communication involves the helper's ability to communicate a constructive response. Carkhuff (1969) cites several studies that suggest "discrimination is a necessary but not sufficient condition for communication." Thus, while a helper may be able to discriminate between a facilitative and nonfacilitative response, he or she may be unable to communicate a facilitative response; but if a helper can communicate a facilitative response, it is assumed he or she can also discriminate as to what is or is not a facilitative response. The implication, then, is to use assessment of communication competence as a measurement of counseling ability. To assess communication competence across the wide range of counseling skills, Gazda (1973) presents a Global Scale, used to rate a response according to how the helpee perceives it. Gazda et al (1977) explain that 1.0 is a harmful response, 2.0 is an ineffective (though not harmful) response, 3.0 is a facilitative response, and 4.0 is an additive response.

This study used the Global Scale (Gazda, 1973) to assess communication competence of the Academy BCT paraprofessionals. The hypothesis was that MC training given the paraprofessionals would be successful. Successful training was defined as training resulting in the paraprofessional attaining a communication competency level of 2.0 or higher on the Global Scale, indicating he or she would at least not be harmful in counseling. Because time and cost constraints limited the study to one which used subjects already chosen for the USAF Academy BCT Paraprofessional Program, a control group was not possible, and the subjects could not be chosen at random; time constraints also prevented multiple administration of the test. Thus, this study used a one-group pretest-posttest design, which involved administering a pretest, conducting the training, then administering a posttest to determine if training made a difference.

Method

Subjects

The subjects for this study were second and third class cadets at the USAF Academy who had been chosen from volunteers for the BCT paraprofessional counselor program. Volunteers were interviewed by a minimum of three experienced counselors who used probing questions requiring self-explanatory, decision-making, and cognitive skills in their evaluation. Trainees were selected on the basis of demonstration of some of the necessary and sufficient counselor qualities, such as positive regard, genuineness, and specificity. Eighty-four cadets selected took the pretest. In addition, one officer acted as a

participant observer, bringing the total number to 85. However, 14 were absent from the training session when the posttest was administered. The number of subjects considered in the analysis for this study, then, is 71.

Instruments

Helpee Stimulus Expressions. Five scenarios, consisting of a situation and statement by a helpee, were read to the subjects. The scenarios represented typical problems paraprofessionals could possibly encounter.

Instructions. Instructions were read to paraprofessional trainees explaining what they were to do for the test and indicating the basis on which their responses would be evaluated.

Global Scale (Gazda, 1973). The Global Scale explained previously was used to to evaluate the subjects' responses to the five scenarios.

Procedure

The pretest was administered to all subjects simultaneously during the first introductory training session. The administrator distributed blank paper to each subject, then read the instructions and scenarios, allowing three minutes between each scenario for the subjects to write their responses. Following approximately 20 hours of MC training, the same procedure was used during the last training session for the posttest.

An experienced Academy counselor assigned each response a rating using Gazda's Global Scale (1973). Tests were identified by the last four numbers of the subject's social security number to eliminate any bias that may have been introduced through knowledge of the subject. Once each response was assigned a rating, the five responses for each test were averaged to give a communication competence score. The data involved two groups: those at or above a certain level of communication competence, and those below. Considering these two groups before and after training, the data were divided into a 2X2 matrix, such that there were four cells. This type of grouping was used with both the 2.0 and 2.6 levels of communication competence as the dividing point for the two groups. To analyze the data, a Chi-Square test was used.

Results

The number of subjects in the two groups, for both conditions, is displayed in the following 2X2 matrices:

I. Level

		below 2.0	2.0 or above
training	before	36	35
	after	13	58

II. Level

		below 2.6	2.6 or above
training	before	65	6
	after	43	28

Applying the Chi-Square test for independence to these data, the following calcuations result:

$$\chi^2_{I} = \frac{142(36x58 - 35x13)^2}{71+71+108+34} = 16.48 \qquad \chi^2_{II} = \frac{142(65x28 - 6x43)^2}{71+71+108+34} = 18.72$$

The table value of χ^2 with one degree of freedom, since (r-1)(c-1)=1, at a level of significance of 0.001 (0.1% level of significance) is 10.83.

Discussion

Interpretation

In the first case (with the division of the groups at the 2.0 level), the resulting value of χ^2 is 16.48, which is greater than the table value of 10.83. Therefore, the four values of each cell of the 2X2 matrix are significantly different than would occur by chance. Though certainty is not possible without a control group, the results imply that paraprofessional training is successful. Thus, 58 of 71 (82%) are competent to counsel without causing harm to the client.

In the second case (with division of the groups at the 2.6 level), the resulting value of χ^2 is 18.72, which is again greater than the table value. Again, the four values of each cell are significantly different than would occur by chance, implying that training did make a difference in bringing trainees to a communication competence level of 2.6 or better. Because a level of 2.6 or better indicates *helpful* responses, these paraprofessional counselors can be facilitative in counseling.

Implications and Recommandations

Considering that training was successful for 82% of the subjects, it seems one can conclude the training program is successful. Certainly, the number of subjects who can counsel at a minimally acceptable level is much higher than if untrained cadre were used. Thus, the program is indeed valuable and necessary. However, even after training, 18% of the subjects rated at a level that indicates communication of possibly harmful responses.

Carkhuff (1969) indicates that both selection and training are factors in effective counseling. Selection based on some of the natural helping qualities of a counselor may be a condition on which the effectiveness of training is contingent. Perhaps this is one reason so many trainees who began training below the 2.0 level increased in communication competence after training--their natural counseling qualities allowed the training to be effective. Although subjects were chosen on the basis of counselor qualities they displayed, the selection process was a subjective one, based on judgments of the counselors interviewing the applicants. A more objective measure of natural helping skills may increase the effectiveness of the training program as a whole. An instrument yielding a measure such as Carkhuff's (1969) Index of Perception could be used to screen volunteers initially to aid in selection of

trainees, and again after training to determine which trainees do not meet minimum standards of counseling competence. Those trainees then would be removed from the position of counselor.

Another recommendation is the use of a control group in a repetition of this study to increase the confidence of a conclusion about the part that training has in increasing communication competence. Additionally, a field survey of the basic cadets would be helpful in determining the effectiveness of the program outside the training environment.

References

Brammer, L. M., The helping relationship. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

Brown, P. R., USAFA Peer-Counseling Program. Paper presented at USAFA, June 1981.

Carkhuff, R. R., Helping and human relations (Vol I, II). New York: Holt, Rinehart and Winston, Inc., 1969.

Egan, G., The skilled helper. Monterey, CA: Brooks/Cole Publishing Co., 1982.

Ford, J. D., Research on Training Counselors and Clinicians. Review of Educational Research, 1979, 49, 87-119.

Gazda, G. M., Human relations development. Boston: Allyn and Bacon, Inc., 1973.

Gazda, G. M.; Asbury, F. R.; Balzer, F. J.; Childers, W. C.; and Walters, R. P., Human relations development (2nd Ed). Boston: Allyn and Bacon, Inc., 1977.

Suicide Gestures of USAF Basic Trainees

Wallace Bloom, Ph.D.

Wilford Hall USAF Medical Center

Introduction

The purpose of this paper is to share our experiences and observations of USAF basic trainees who made suicide gestures. Lackland Air Force Base is the sole location for basic training of all USAF recruits, and on June 1, 1975, an experimental Air Force Medical Evaluation Test (AFMET) Program (Bloom 1977) was started there to facilitate early identification of unsuitable trainees. In AFMET Phase I, initial screening was by a computer scored test. On the second day of training, all recruits take a 50 item, true-false Historical Orientation Inventory with four additional demographic items. On the basis of twenty-eight (28) items previously correlated with failures to complete enlistments, the 6½% most vulnerable are identified for further evaluations. At Phase II (Interviews and Tests) appraisals are noted on Standard Reports of Interview (SRI)(Bloom 1981) and Bloom Sentence Surveys (BSCS) are taken (Bloom 1979, 1980a). About two of each 6½ per hundred are referred to AFMET Phase III (Clinical Assessment), the others are cleared for return to duty. Usually between the 8th to 12th day of training, officer clinicians conduct diagnostic interviews and use a variety of individually selected psychological tests. About 16% of Phase II selectees have been recommended for administrative separations. During the eight years of the AFMET Program, there has not been any death by suicide of a basic trainee. The statistical expectancy was 20 suicides or two and a half each year. This is based on similiar age civilian population, and considered only the minimum 10,000 trainees here at one time rather than the 70,000 to 80,000 trained each year.

Background

For reference purposes, there were 1446 suicides of USAF personnel during the twenty year period ending November 2, 1979, average 72 per year. This was significantly less, about one half of a similar age civilian population. After the experimental one year research project was completed in June 1976 testing was suspended. The Air Staff reviewed the experimental results and decided that the AFMET project would be made operational. During that summer while testing had been temporarily stopped, one lethal suicide occurred. Testing was restarted immediately one week before previously scheduled date. There has not been a suicide since then or during the earlier AFMET year.

During fiscal years 1979 and 1980 there were 55 to 65 trainee suicide gestures or attempts respectively. By 1982 and 1983, these had been reduced to 24 to 30. Last year there were 16 cutters, 11 pill takers, and 3 jumpers and almost all were [illegible]arily attention-getters of low lethality. At least three types of suicide [illegible]res have been tentatively identified. First, those in the first five days [illegible]ing by individuals who prior to reporting, wished to cancel delayed en-[illegible] because in the meanwhile they got a better job, a school admission, a [illegible]ale up with their family. Second, were individuals who felt overwhelmed [illegible]ually between the 6th to 15th day. The third group, after the 24th [illegible] who felt unable to face four years of service or had been selected [illegible]raining or locations they did not desire. Actions of the first group

were based on the past, of the second on the present, and of the third, their perception of the future.

Prevention

Although the AFMET project selects only 6½% of all trainees for Phase II, 18 to 21% of those who made suicide gestures in 1979 and 1980 were Phase II selectees. We therefore have been vigilant for significant clues during Phase II interviews and tests particularly regarding item 16 (serious suicidal ideations or previous attempts) on the Standard Report of Interview (SRI). Approximately 15% of trainees report these. On the Bloom Sentence Completion Surveys, we found that individuals who have been contemplating suicide at some specific future time have difficulties responding to items beyond that time on items 15, 31 and 39. A psychological brick wall barrier seems to have been erected at that date so they may omit answering, for example, No. 15 "Ten years from now" or evade by responding "I just live from day to day," or "I don't know." Responses to item 12 often seemed sensitive to suicidal thoughts and was the fourth indicator.(Bloom 1980b) When two out of four replies were other than positive, the patient was usually referred for further clinical assessment, Phase III.

Studies of actual suicidal incidents and gestures in 1978 and 1979 revealed a higher frequency between the 5th through 15th day of training. The school commander then instituted a four hour recreation break on the 6th day to relieve the pressures. The frequency of gestures were reduced. Other beneficial activities included the yearly psychological retesting of all instructors. This identified some who were starting to burnout or had developed other emotional problems for counseling or assignment changes. The trainees benefited as healthy instructors create less stress for them and are more likely to spot trainees needing mental health referrals. The ready accessibility of satellite mental health clinics in or close to trainee barracks, facilitated communication as the mental health staff worked within the training situation rather than apart from it. From frequent contact with clinic staff, instructors became more adapt at spotting early behavioral clues to recruit'sproblems and made appropriate and prompt referrals. Squadron escorts were provided for edgy patients and less acting out occured to, from, and within the clinics.

Intervention

In one instance when a recruit went out on a third floor window ledge, a staff member of the mental health clinic, talked him back to safety. Another, who perched on a third floor fire escape railing, was brought back after more than 2 hours talking from a chaplain assisted by a psychologist who had rushed over as soon as the clinic was notified. Action has been taken to insure prompt notification of future crisis situation when mental health action is appropriate.

Postvention

Need for follow-up after suicide attempt or gesture was demonstrated when the previously cited ledge crawler was diagnosed at the hospital as having Agoraphobia and recommended for discharge. He came back to the clinic to take a battery of tests as we were collecting data for suicide research. It was further found he had Oc' `ophobia (fear of crowds), had panicked in the crowded dorm and initially went out the window for needed space. He might have panicked again during some of the formations and crowded briefing of outprocessing. We contacted the Casual Squadron Commander and arranged for this airman to follow behind formations and sit alone in back of classrooms as he had done since his junior high school. Preliminary arrangements were made for him to go to stress management and other therapies on his return home.

As needed, trainees in the casual squadron being processed for discharge have returned to the clinic for therapy, often on a daily basis. The clinic has phoned parents of suicide attempters to learn if there had been previous attempts, threats or suicides of acquaintances. With the attempter's permission, inquiries have been made concerning past or future therapy.

Conclusions

Suicides and suicide gestures do occur within the Air Force, though less frequently than in like civilian populations. Many incidents can be averted by preventive programs of mental health clinics and cooperative actions of line organizations. Further organized research and data collection into a central bank of psychological information on each incident, (to include precipitating factors) is recommended. Activites of mental health clinics can change to meet emerging Air Force needs(Bloom 1983). In 1983, only one attempter was an AFMET Phase II selectee, in contrast to the ten to twelve in 1979 and 1980. Intensive training of interviewers for early identification of trainees who are above normal suicide resks have paid off. Individuals so identified, get priority intensive assessments, may be defused by brief therapies, or recommended for immediate removal from training and discharge. It is not that AFMET interviewers are seeing fewer high risk individuals but we are dealing with them more effectively.

References

Bloom, W., Air Force Medical Evaluation Test. Medical Service Digest, USAF 1977, 28 (2) 17-20.

Bloom, W., Bloom Sentence Completion Survey, Student: Instruction Manual. Chicage, Illinois. Stoelting Co. 1979.

Bloom, W., Bloom Sentence Completion Survey, Adult: Instruction Manual. Chicago, Illinois. Stoelting Co. 1982.

Bloom, W., Values of Idiosyncratic Responses in a Mass Psychological Screening Program. Society for Personality Assessment, Annual Convention, Tampa, Florida. March 15, 1980b.

Bloom, W., Assessment of Patients, Standardized Report of Interview(SRI). Military Medicine 1981, (Sep) 146, (9) 639-641.

Bloom, W., Changes Made, Lessons Learned After Mental Health Screening, Military Medicine 1983 (Nov) 148, 889-890.

AD-P003 246

Factor Stability and Construct Validation of Yukl's MBS (Managerial Behavior Survey) for Military Leadership

David D. Van Fleet
Texas A&M University

Gary A. Yukl
State University of New York at Albany

Abstract

Yukl's new, complex taxonomy of leader behavior has been examined in a variety of contexts including that of military leadership. As with all new concepts and instruments, however, replication is necessary. This paper presents an effort to establish, through data collected at two points in time for the same organization, that Yukl's MBS does display both stability in the factors obtained for these two time periods and that those factors do indeed represent the 23 dimensions covered by the instrument. While the sample used is a military cadet sample rather than one of regular military officers, these results are sufficiently encouraging to suggest that more research should be performed using this new taxonomy. Directions that such research could take are discussed in light of these and other results.

Introduction

Much previous research on leadership in general and military leadership in particular has been limited for at least two major reasons. First, an overly simplistic conceptualization of leader behavior has been all too frequently used in that research. That usual dichotomy of task- and maintenance-oriented (structure and consideration, production-centered and people-centered, etc.), while highly generalizable, is hardly very useful for the development of training and developmental materials or in selection and evaluation. Second, the overwhelming strategy in much of the existing research is to use a single point in time (cross-sectional analysis) and a single method (usually the questionnaire-correlational method). While certainly convenient and useful in the early stages of research, this strategy must give way to improved ones if we are to press the boundaries of knowledge further in this area.

Yukl (1981) has presented a new taxonomy of effective leader behavior designed to overcome the first limitation noted above. His taxonomy is at an

intermediate level of analysis, that is, at a more specific, detailed level than the usual task- and maintenance-oriented level and yet not so detailed and specific as individual job analysis would be. The application of that taxonomy to military leadership has already been successfully demonstrated in a multi-method study (Yukl and Van Fleet, 1982). Thus, a first step has been taken to overcome the limitations of the past in order to make substantial progress for the future.

Purpose

The purpose of this paper is to present some evidence regarding the factor stability and construct validity of the new taxonomy for military leadership. In doing this, a second step will be taken in that a comparative analysis of data from two points in time will be performed rather than merely using cross-sectional data as in the past.

Method

The sample utilized in this study consisted of members of the Texas A&M University Corps of Cadets. The Corps is organized into military units; its members wear uniforms while on campus and participate in frequent drill, formations, and military ceremonies; and many of the members of the Corps are in the ROTC programs of the several services. In fact, over 14,000 officers have been commissioned through the Corps, and more that 100 of them have attained general officer rank.

The method consisted of administering Yukl's Managerial Behavior Survey (MBS) at two points in time. The first administration was after the first full semester (early in the Spring Semester) while the second administration was near the end of the second full semester (late in the Spring Semester). This assured that those asked to provide information about the leaders would have had ample opportunity to interact with and observe those leaders and that the two points in time were reasonably separated (about three months). Participation was voluntary so that the second administration resulted in a lower participation rate than did the first (597 observations for Time One and 261 for Time Two).

A factor analysis was performed for each time period. From these results, then, two analyses were possible. First, the number of factors obtained, the percent of variance accounted for by those fators, and the specific content of those factors can be compared to examine the factor stability of the MBS. Second, the specific content of the factors can be compared to the categories in the taxonomy as coded in the scoring of the questionnaire to see if the factor analysis is indeed extracting factors which correspond to those categories.

Results

As is obvious from Figure 1, both factor stability and construct validation were demonstrated for this sample. The number of factors obtained is nearly identical; the percent of variation accounted for is remarkably consistent; and the content of the factors is very similar. All of this assures that the MBS instrument does possess factor stability for this sample. Coupled with previous results (Yukl, 1982), the evidence is very strong that

this instrument possesses factor stability as well as good reliability (see Yukl and Nemeroff, 1979).

Insert Figure 1 About Here

The content of the factors relative to the categories from the taxonomy was examined by comparing the five items which were scored to yield the categories from the taxonomy with the highest loaded five to ten (out of 115) items from the factor analysis. For the first time period, where the sample size was much better for performing factor analysis, 21 of the factors have, as the five highest loaded items, the five which define a particular category. That number is increased to 22 of the 23 when the top 10 highest loaded items are examined. One other factor had four of the five loaded on it. For the second time period the results are also quite strong although not so overwhelming as with the first time period. These results, then, strongly suggest that the MBS displays construct validity in that extracted factors correspond extremely well to the categories from the taxonomy represented by that instrument.

Conclusions

The data presented here strongly suggest that Yukl's Managerial Behavioral Survey (MBS) possesses both factor stability and construct validity as well as previously demonstrated reliability. This means that the MBS can, indeed, be used with military samples to extend our knowledge about effective leader behavior. Such extensions will be even more meaningful, of course, if the research strategies used go beyond single method, single time strategies. If future research will use this more realistic, complex taxonomy in more useful and more complex research strategies, much can be learned about leadership in general and military leadership in particular which can be particularly useful in selection, evaluation, training, and development of future military officers.

References

Yukl, G. A. Leadership in organizations. Englewood Cliffs, New Jersey: Prentice-Hall, 1981.

Yukl, G. A. Innovations in research on leadership behavior. Paper presented at the Eastern Academy of Management, Baltimore, Maryland, May, 1982.

Yukl, G. A., & Nemeroff, W. F. Identification and measurement of specific categories of leadership behavior: A progress report. In J. G. Hunt and L. L. Larson (Eds.) Crosscurrents in leadership. Carbondale, Illinois: Southern Illinois University Press, 1979.

Yukl, G. A., & Van Fleet, D. D. Cross-sectional, multi-method research on military leader effectiveness. Organizational Behavior and Human Performance. 1982, 30, 87-108.

Figure 1

Factor Stability and Construct Validation of MBS Categories

Scale	Time One: Number of Items Defining Scale Loaded Among Highest on Factor (maximum is five) — Highest Five	Time One — Highest Ten	Time Two: Number of Items Defining Scale Loaded Among Highest on Factor (maximum is five) — Highest Five	Time Two — Highest Ten
Showing Consideration	5	5	5	5
Providing Praise and Recognition	5	5	5	5
Training-Coaching	5	5	5	5
Disseminating Information	5	5	5	5
Encouraging Decision Participation	5	5	5	5
Delegating	5	5	5	5
Innovation	5	5	5	5
Facilitating the Work	5	5	5	5
Monitoring the Environment	5	5	5	5
Representing the unit	5	5	5	5
Managing Conflict	5	5	5	5
Emphasizing Performance	5	5	4	5
Inspiring Subordinates	5	5	4	5
Goal Setting	5	5	4	5
Planning	5	5	4	5
Criticizing	5	5	4	5
Career Counseling	4	5	5	5
Problem Solving	5	5	4	4
Clarifying Work Roles	5	5	3	4
Administering Discipline	5	5	3	4
Facilitating Cooperation and Teamwork	5	5	2	2
Monitoring Operations	5	5	1	1
Structuring Reward Contingencies	4	4	4	4

	Time One	Time Two
Percent of Variance Accounted for by:		
23 factors corresponding to scales	70.78	78.74
all factors extracted	75.54	83.60
Total Number of Factors Obtained	25	26
Number of Observations	597	261

Exploring the Interaction of the Vroom/Yetton Model and Leadership Style (LPC) as it Predicts Performance

by
Major Robert M. McDannell
and
Captain Scott C. Marcy
United States Military Academy

Abstract

The Vroom/Yetton (1973) model prescribes the decision making method most appropriate for a given situation. The decision should be within the area of authority of the manager and have some forseeable effect on the organization and subordinates. Since it is a prescriptive model, leader behavior that is consistent with it should occur more for effective leaders than for ineffective leaders. The purpose of this research is to explore the relationship between leaders' decision-making behavior on the Vroom/Yetton problem set and performance in an actual organizational setting. Additionally, leadership style as measured by Fiedler's Least Preferred Coworker (LPC) scale is tested as a personality variable that may moderate the problem set to actual leader performance relationship. A group of 98 third year cadets at the U.S. Military Academy were administered a military version of the Vroom/Yetton thirty problem set and Fiedler's LPC scale. Military development ratings made by their Tactical Officers were used as measures of their performance. Results suggest that information resulting from their performance on the problem set is related to their performance evaluations in an actual setting. Additionally, the magnitude of the relationship is larger for low LPC cadets and disappears for high LPC cadets.

<u>Background</u>. The Vroom/Yetton model prescribes a feasible set of decision making styles for a leader based on the nature of the decision making situation (Vroom & Yetton, 1973). There are five decision making styles which vary on the amount of subordinate participation ranging on a single dimension from fully autocratic to fully participative. A leader determines what style is appropriate based on the presence or absence of seven problem attributes that are likely to influence the effectiveness of the decision making styles. These attributes include: quality requirement, leader information, problem structure, necessity of subordinate acceptance, leader certainty of selling autocratic decisions, subordinate sharing of organizational goals and likelihood of conflict over preferred solutions. Work by Vroom and Yetton (1973) and Vroom (1976) provided partial support for the model but were challenged based on experimental design shortcomings. Vroom and Jago (1978) examined the model's validity in a well designed study in which they first trained 96 managers in decision processes and the attributes of quality and acceptance of decisions. Next the managers described two decisions that they had made: one that seemed to be successful and one that did not. They were then trained on the theoretical

problem attributes of the model and determined whether these attributes were or were not present in the situations they had previously described. The results of this study provided substantial support for the validity of the model as it applied to actual decision making situations.

However, as Field (1979) pointed out, a major threat to the external validity still exists since the model was developed and tested using self-report techniques. He found that self-reported decisions could be biased toward important, unambiguous or short-term problems, because reporting such decisions would be more meaningful or easier for the manager. Generalizing the model to less important, long-term or problems where attributes are not as clear could prove difficult. To avoid the bias inherent in self-reporting behavior, a reasonable starting point would be to explore the predictive power of the Vroom/Yetton model. If the model indeed measures appropriate decision making behavior than it would seem that those who perform better on the problem set should in general be more effective leaders. Additionally, both raters and subjects should be unaware of the propositions of the model, which is the case in the present study.

> H1: Leaders who are not trained in the Vroom/Yetton model and perform better on the Vroom/Yetton problem set will perform better in actual organizational settings.

The principal issue of this research is to explore the effect of leadership style on the relationship in H1 above. According to Fiedler (1967), two people of different leadership styles as measured by LPC will naturally behave differently in identical situations. If one accepts Fiedler's proposition, then it would seem that leadership style may affect decision-making behavior on the Vroom/Yetton problem set as well as actual performance. If this is true, then the relationship between problem set behavior and actual behavior may be partially suppressed or at least affected by leadership style. At this point in our research, the likely direction of this effect is not hypothesized. For now, as an exploratory exercise, a simpler hypothesis is in order.

> H2: Leadership style as measured by LPC moderates the relationship between leader behavior on the Vroom/Yetton problem set and performance in actual organizational settings.

Method. Data were gathered from 98 third year cadets then taking a required course entitled Military Leadership. As part of learning exercises, cadets completed both Fiedler's LPC scale and a military version of the Vroom/Yetton problem set, which is theoretically identical to previous civilian versions in terms of problem attributes. From cadet responses, the data gathered include frequency of: use of each decision style, rule violations, agreement with the feasible set, agreement with the most and least participative style of the feasible set (Model A/B), mean level of participation and leadership style (LPC).

For a behavioral measure of performance we use the cadets' military development ratings done by their company tactical officers. The items of the instrument describe behaviors that reflect military development. Cadet scores on these scales are primary determinants of the level of appointment to leadership positions as seniors. This is important as most of the cadets were not actually in leadership positions when these ratings were done. Therefore, this measure may be said to be related to leadership performance but does not constitute direct measurement of it. The behavioral dimensions of the scale are: task structuring and management, interpersonal relations, compliance with organizational expectations, intellectual application and growth, personal professional ethical behavior and performance oriented development. Each dimension is tapped by three items, each of which has three behaviors yielding item scores that range from one to five.

Results of analyses. Frequency of agreement with the feasible set is used as a single measure of problem set performance. We created two groups by splitting the sample at the median of this variable. Comparison of means on the six dimensions of performance yield no significant differences. However, splitting the sample by the amount of participation allowed in the problem set choices yield significant results for compliance with organizational expectations and performance oriented development. In the interest of examining main effects of feasible set agreement and participation, a global measure of performance was computed by taking the average of all six performance dimensions. The t-test yields significantly higher performance for the high participation group. A similar comparison by splitting on the median of feasible set agreement again yielded no significant difference.

The most interesting results acheived are the differential effects of participation. There is a main effect for participation that is enhanced for the low LPC group and is absent from the high LPC group.

Mean Scores on Global Performance Moderated by Participation on Vroom/Yetton Problems and LPC Score

	Full Sample	Low LPC	High LPC
Low Participation	4.22 n=44	4.14 n=20	4.29 n=24
High Participation	4.43 n=46	4.52 n=29	4.28 n=17
t =	2.11 $p<.05$	2.85 $p<.01$	.03 ns

Additional evidence of the effects of leadership style is apparent from the multiple correlation coefficients between the Vroom/Yetton predictors (frequency of rule violations, agreement with the feasible set, etc.) and each of the six dimensions of performance (task structuring, interpersonal relations, etc.). The average multiple R square between the predictors and performance for the entire sample average .17 compared to .37 for low LPC and .39 for high LPC. The reason that the two sub-populations both have stronger relationship is because the sign of the correlation coefficients change consistently between the low and high LPC groups. This will be more closely examined in future research.

Discussion. The relationship between actual peformance and the Vroom/Yetton problem set behavior is weak. Although all of the Vroom/Yetton predictors account for roughly 17 percent of actual performance, the single predictor of agreement with the feasible set accounts for virtually nothing. There are two plausible explanations for this that are not the fault of the model. Cadets, although they are students of leadership, have little experience as leaders and managers. Their performance, as measured by agreement with the feasible set, on the Vroom/Yetton problem set is worse than that of trained managers. Cadets' mean agreement is 60% while Vroom and Yetton (1973) report 69% and Vroom and Jago (1978) 65% for professional managers. Also, cadets tend to be more autocratic than most populations as shown from their population mean participation score (3.79) compared to college students (5.61) and managers (4.68) (Vroom & Yetton, 1973). Cadets are even more autocratic than the Vroom/Yetton autocratic model (4.17). Cadets are more likely to choose from outside the feasible set erring toward the autocratic. This tendency may result in a lack of variance that weakens the true predictive power of the model.

Simple participation is the stronger predictor and is further strengthened by leadership style. This relationship between participation and performance supports earlier propositions of the positive effects of participation by Blake and Mouton (1964) which is a more simplistic approach. Of importance is the idea that LPC or task motivated cadets tend to be more favorably evaluated by their raters if they were more participative on the problem set even if they violated decision making rules in the participative direction. If the climate of the cadet companies could be classified as high situational control by Fiedler's criteria (good relationships and high structure) then it is theoretically consistent that high socioemotional behavior which is characteristic of the low LPC here, would be more effective. In this research it is interesting that a person who is task motivated and demonstrates a propensity toward participation actually performs better. Based on the above assumption of situational control, this provides support for Fiedler's contingency propositions. That is to say that the task motivated leaders are participative in the appropriate situations. It is not clear at this point what the differential effects of decision training for low and high LPC populations may be.

Practically speaking, these results suggest that cadets who demonstrate a willingness to share decision making with subordinates are perceived more favorably by their tactical officers and moreso if they are task motivated. Furthermore, the idea of participation in leadership and decision making, if related to performance as demonstrated here, lends credence to the teaching of this to students of leadership in military settings. This may serve to further enlighten the student to share information and power rather than to hoard it.

The interaction of decision making and leadership style warrants more study. It appears that both theories might be enhanced if each incorporated components of the other. If the predictive power of the Vroom/Yetton model can be enhanced by LPC then perhaps it should be considered in the training or prescriptive nature of the model. A simple 2 X 2 field experiment wherein one group received Vroom/Yetton decision making training, another receives Fiedler's Leadermatch training, a third gets both and the fourth or control group gets no training would constitute a fairly rigorous test of the external validity of these models.

As mentioned above, cadets at West Point gain some understanding of both the Vroom/Yetton and Fiedler models. However, they are not specifically "trained" on either. Leadermatch is well validated as it improves actual leader performance. Vroom/Yetton is perhaps less so. If the interaction of the two theories (which apparently exists) could be more clearly defined, then their value as leader training devices could be increased.

Vroom/Yetton prescribes a conscious selection of style based on the situation while Fiedler prescribes altering the situation based on personality. Both theories include issues of task structure the degree to which subordinates support the leader and organization, and the amount of personal and administrative influence the leader has on subordinates. Further research may lead to prescriptions that are both more explicit and more effective in terms of organizational outcomes. For example, Vroom/Yetton feasible sets may be further delimited depending on leadership style or; decision making styles could be specified for a leader depending upon his situational control. The prospect of a coupling of the theories is truly intriguing and represents an area of inquiry important in both theory and application.

REFERENCES

Blake, R., Mouton, J. The managerial grid. Houston, Texas, 1964.

Fiedler, F.E. A theory of leadership effectiveness. New York: McGraw-Hill, 1967.

Field, R.H.G. A critique of the Vroom-Yetton contingency model of leadership behavior. Academy of Management Review. 1979, Vol. 4, No. 2, 249-257.

Vroom, V.H. Can leaders learn to lead? Organizational Dynamics, Winter 1976, pp. 17-28.

Vroom, V.H., Jago, A.G. On their validity of the Vroom/Yetton model. Journal of Applied Psychology, 1978, 63, 151-162.

Vroom, V.H., Yetton, P.W. Leadership and decision making. Pittsburgh, Pa. University of Pittsburgh Press, 1973.

SUBORDINATE PERCEPTIONS OF CONTINGENT LEADERS: DO FOLLOWERS ACCEPT OUR THEORIES?

Jack H. Cage

United States Military Academy

Abstract

Two studies were conducted to investigate assumptions made by contingent leadership theorists about subordinates. Paper-and-pencil instruments provided situations in which a leader followed or failed to follow the prescription of Vroom's contingency theory. 393 subjects were asked to take the role of subordinate, evaluate the decision-making process, offer prognosis about the outcomes, and assess the leader. Three groups of subjects were identified by their responses. Only one of three groups accepted the leader when s/he acted according to the prescription of contingency theory. The groups were differentiated by characteristics such as income, occupational prestige, and number of subordinates. The findings suggest that individuals, while acting as subordinates, may fail to accept the prescription of contingency theory.

Introduction

Background

In the long, exhaustive history of leadership research, investigators have viewed leaders' traits, behaviors, and the interaction of leaders with situational factors to determine those stylistic features that most directly affect their success. This last group, the contingency theorists, has risen to prominence over the last twenty years due to studies that appear to explain some of the variance involved in determining leaders' success. Contingency theorists assume that a leader interacts with certain situational factors to affect a range of outcomes. One major theorist, Victor Vroom, focuses on the specific problem faced by the leader and on the degree to which the leader includes subordinates in making highly accepted, high quality decisions. Vroom hypothesizes that a leader's behavior and the specific situational factors that s/he faces interact to affect his/her success.

Although the work in contingency leadership theory has been productive and has significant practical and theoretical implications, one major consideration has been overlooked: How do

subordinates of theoretically correct contingent leaders view their leaders' actions? That is, do individuals while acting as subordinates accept the prescription of contingent leadership theory? Do they, for example, expect a leader to be autocratic in one situation and participative in another, or do they prefer more consistency from their leaders? Heilman and Hornstein (1981) suggest that contingent leadership theorists make five implicit assumptions about subordinates: (a) subordinates have an implicit theory of leadership that guides their evaluation of the leader; (b) subordinates assume an orthogonal independence between the leader's responses in different situations; (c) subordinates are assumed to have the same information as the leader and process it in the same way; (d) subordinates are assumed to use the same criteria for evaluating all leaders regardless of personal characteristics; and (e) subordinates apply the same rules to leader behavior uninfluenced by the organizational context surrounding the encounter.

It may or may not be true that contingent theorists' implicit assumptions about subordinates are fully justified. If invalid, however, these beliefs may cloud the actual interactions between a senior and his/her subordinate. Because of these important questions and the implications that they hold for contingency theory in general, two studies were conducted to investigate these assumptions, and to determine just how subordinates evaluate theoretically contingent leaders. The first study was designed to investigate the assumption that subordinates hold an implicit theory of leadership. The second study (Cage, 1982) was conducted to study the assumptions of interchangeable leaders and context, to include the assumption of an implicit theory. Both studies, however, used similar procedures and the results reported here are based on a similar outcome.

Methods

In both studies, paper-and-pencil instruments based on earlier work by Vroom provided situations in which a leader followed or failed to follow the prescription of Vroom's contingency theory. In all cases, subordinates were asked to assume the role of subordinate.

Because of the specific assumptions under investigation, the first study provided a random grouping of situations to 25 randomly selected subjects, each of whom read six stories or cases. These cases were divided into two groups; the autocratic decision-making process was theoretically prescribed for three stories while the participative process was prescribed for the other three. The manager described in each case used one of three processes (autocratic, consultative, or participative). The procedure provides two cells in which the prescribed and actual behaviors co-occur. In other words, the leaders described in the first column are presented with certain situations that would

theoretically require them to apply an autocratic decision. Thus, a 3 X 2 within-subjects design presented the prescribed behavior and the leader's behavior as the independent variables.

The second study, instead of within-subjects, employed a factorial design. The situations used in the first study were modified to include information about the leaders' personal characteristics as well as the context in which they operated. Additionally, the consultative option was removed; the leader could apply either an autocratic or a participative response. 368 male MBA candidates each responded to one of these stories.

Both studies used 21 bipolar adjectives as dependent variables. Subjects were asked to evaluate the decision-making process, offer prognosis about its outcomes, and assess the leader.

Results

If Vroom's model was a correct representation of reality, we would expect to find that subjects rated theoretically appropriate leaders and their decisions as highly effective. For example, leaders who behaved autocratically would be seen as highly competent when faced with an autocratic situation, while, in the same situation, a leader who behaved participatively would be rated as less effective. The results of the studies demonstrate that these expected findings were not consistent, and that many subjects rejected theoretically appropriate contingent leaders especially when these leaders were autocratic. These findings were not consistent with expectations, and were the basis for subsequent analysis reported here.

While investigating the failure of subjects to endorse theoretically appropriate autocratic leaders, subpopulation differences were identified. Three different groups, initially identified by their patterns of responses, were differentiated among subjects who participated in the two studies. One dependent measure, ratings of the leader's competence, along with certain associated demographic variables are presented in Table 1; the data are representative of the patterns found with other measures of decision/leader effectiveness. Because of the differences in the designs of the two studies and in the numbers of subjects, the data can not be compared statistically. They are, however, useful for suggesting trends worthy of future investigation.

As Table 1 shows, group #1, identified in the results of the first study, consisted of subjects who prefer participation regardless of situational constraints. The group was characterized by its 'lower' standing on several demographic variables: income, occupational prestige, and number of subordinates. Group #2 included subjects from the second study who preferred participation in appropriate situations yet rejected autocratic behavior. The

group was characterized by its 'higher' standing on demographic variables: income, occupational prestige, and number of subordinates. Group #3, identified in both studies, consisted of those subjects who rated theoretically correct contingent leaders as most effective. This group was intermediate between the other two groups on demographic measures.

Table 1

Means and Standard Deviations for Three Patterns of Ratings:

Leader's Competence

	SUBJECTS' RESPONSE PATTERNS		
	Reject Autocratic "Lower" (#1)	Contingent Group "Middle" (#3)	Accept Appropriate Participation "Higher" (#2)
INCOME ($K)	20.2 (3.6)	27.2 (4.8)	42.3 (3.3)
PRESTIGE	28.7 (7.4)	49.2 (3.6)	56.5 (2.5)
AGE	32.6 (4.6)	31.7 (2.5)	33.4 (4.2)
NUMBER OF SUBORDINATES	.2 (.2)	2.5 (2.8)	6.5 (3.3)

NOTES: 1. N for each variable in Column #1 = 16
2. N for each variable in Column #2 = 58
3. N for each variable in Column #3 = 66
4. Prestige ratings have a range of 14 to 85.

Discussion

One interesting argument can explain the relationship between the subjects' responses and their standing on demographic measures. These groups' responses may reflect their vastly different reasons for working. Members of the 'lower' group may work merely for security; they are not 'followers'--active and willing participants--but subordinates based on a structural relationship with a superior. These people achieve a measure of security with participation that they would not have with another, possibly biased leader/manager. In other words, a participative leader would serve to increase their satisfaction. The assumptions and interests of the 'middle' group are different. The members of this cohort are more powerful than those in the first. They, for example, are better educated, have higher levels of occupational prestige and income. Their own interests may be tied to the organization and its success; the members of this group, who haven't quite 'made it' to the degree that the third group has, are aspiring 'followers.' They view the most effective processes as personally beneficial since their goals and those of the organization are linked. These individuals can tolerate an autocratic boss since to do so is in their own best interests. The same can not be said for the 'upper' group. Members of this

cluster are firmly entrenched in the organization. They are more powerful, more creative, of higher status, and have a different set of experiences to guide their expectations. They do not expect to be led; they expect to lead, or, at a minimum, to participate. Autocratic leadership behavior is never perceived as effective by this group; they have crucial information and are concerned with less structured, long-range problems facilitated by participative processes. These people are not indiscriminately endorsing participative behavior, but they always disapproved of behavior that is autocratic.

These findings suggest that individuals, while acting as subordinates, may fail to accept the prescription of contingency theory. Furthermore, they suggest that subordinates' demographic characteristics are associated with their differing expectations and implicit theories of leadership. Leaders must at least consider these characteristics before deciding upon the leadership process to employ. For as these findings suggest, subordinates may not wholeheartedly 'buy into' commonly accepted theories of contingent leadership.

References

Cage, J.H. "Subordinate Perception and Evaluation of Leaders Who Differ on a Personal Characteristic: Are Leaders Interchangeable?" Unpublished Doctoral Dissertation, Columbia University, 1982.

Heilman, M.E., and Hornstein, H.A. What is Effective Leadership Behavior: The Subordinate Perspective. Unpublished manuscript. 1981.

Vroom, V.H. and Yetton, P.W. Leadership and Decision-Making. Pittsburgh: University of Pittsburgh Press, 1973.

To the Wilderness and Beyond: The Application of a Model for Transformal Change [1]

Robert C. Ginnett
Yale University (AFIT/CI)

Much of the ongoing management and development in modern organizations could be considered transactional in nature, and as long as the organization is operating reasonably effectively and efficiently, this style may be most appropriate. However, there may also be organizations in which transformal leadership or development is needed. A model for transformal change was applied in such an organization with qualitatively successful outcomes. The model and its application are described as well as a discussion of the dilemma of evaluation for efforts of this nature.

Background

James MacGregor Burns described two fundamentally different kinds of leaders: the transactional leader and the transformal leader (1978). The former is most efficacious at performing the tasks of management of the internal functions of the organization. If a maritime analogy could be drawn, the transactional leader would be like the navigator, who when given a destination, would be responsible for plotting the course to the goal, and detecting and correcting deviations from that course. Continuing the analogy but with another focus, one might consider the person who envisions and communicates the hoped-for destination as the transformal leader. This distinction between transacting and transforming has also been drawn by Zaleznik although he has characterized the two types as managers and leaders respectively (1977). While the curricula of the nation's management schools are replete with the essence of transactional management, one is hard pressed to even describe the mechanism of transformal leadership, let alone detail its processes. This aspect closely parallels the "intuitive" characteristics reported in Mintzberg's seminal emperical study of successful executives (1975).

Suggesting that transformal leadership is largely intuitive leaves us with little to offer the leaders of today. But if we focus instead upon the functions required for organizational effectiveness, then the entire burden for transformation may not fall solely upon the single legitimate leader but may arise from any number of sources. If given the opportunity, there may be employees at various levels who might "create a vision" of a better organization.

In any event, the shift to a functional orientation does not resolve our difficulty but merely changes its focus. We must now

---

1. The views expressed in this article are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

ask, "how can the transformal process be made available in an organization which has been geared to operate transactionally?" The opportunity to move this question from academe to the workplace came quite serendipitously to this author in an offer to participate in a consulting project.

The Organization

The client under study was a large human services organization (HSO) in New England, structured with a commissioner, two deputy commissioners in charge of line (service) and staff (support) functions respectively and a variety of subordinate managers of regional and institutional services. The problems in this HSO far exceeded those characterized as typical under Domain Theory (Kouzes & Mico, 1979). The internal problems had reached such proportions that a spontaneously generated committee of concerned managers had determined that the department could not perform its function without a reorientation and improved team cooperation. The downward "lack of trust-lack of communication" spiral was evidenced throughout the department but perhaps no better exemplified than at the top where the two deputy commissioners and their staffs had not had a face-to-face meeting in the previous two years! In this atmosphere, the management development committee and the consulting team set out to transform the organization. Fortunately, there was a shared awareness that the change would not be easy, nor would it be quick. The planning stage alone consumed four months.

A Model and Process for Transformal Change

The consultants grounded their role in a conceptual model [2] presented in simplified form in Figure 1.

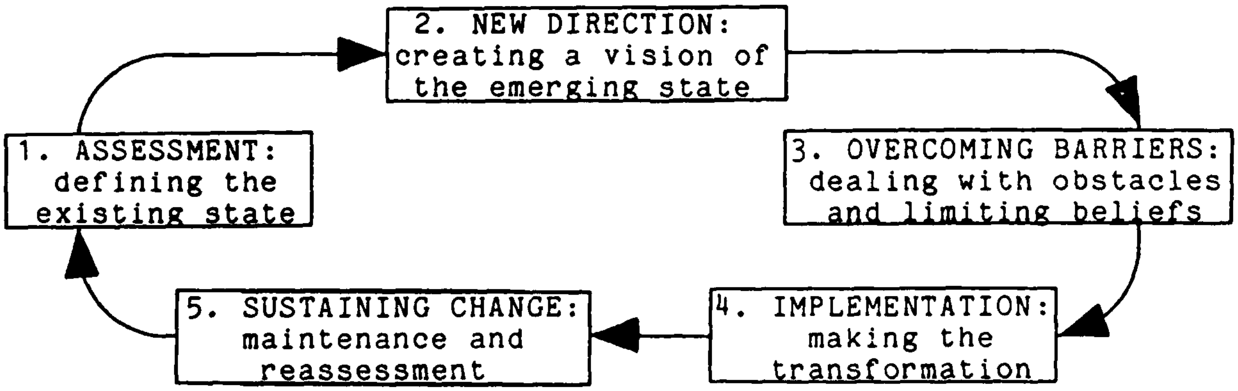


Figure 1. A Model for Transformal Change

The first of the five elements was designed to surface the varied perceptions of the existing state typical in HSOs (Kouzes & Mico, 1979). The development of a collective vision in the second step

---

2. Both the model and the opportunity were provided by Dennis N.T. Perkins, to whom I am grateful.

permits the group to then deal with barriers in a new light in step three. Although the first three steps seem necessary for transformal change, they are hardly sufficient. Step four requires that roles, responsibilities and coordination be made explicit and step five is designed to operationalize and continue the process back in the organizational setting. Obviously, the application of this model is more complicated and requires more explanation than space here permits. However, the process was designed to occur in three phases.

<u>The Wilderness Experience</u>. In this organization, the history of mistrust made "unfreezing" (Lewin, 1951) critical; thus a somewhat radical technique was employed. This phase, which was designed to demonstrate experientially the need for teamwork, was conducted literally in the wilderness. Three 15 person teams (plus one consultant each) were constructed based upon a stratified random design to balance gender and organizational status and level. After an initial team building process, each of the three teams of managers were taken into a wilderness area by qualified staff. The participants (age range: 35-65) were required to survive in this rugged environment as a team for three days. In addition to subsistance functions, tasks included approximately 25 miles of hiking over mountainous terrain, stream crossings, and technical rock climbs. Although somewhat similar to the "Outward Bound" experience for managers, this experiential exercise had the advantage that all participants were from the same organization.

<u>The Application Session</u>. An even more significant difference here than in Outward Bound was the integrated follow through. Held in an off-site location approximately two weeks after the Wilderness Experience, this two-day session was designed to first reintegrate the three teams. (Not unlike the classic Robbers Cave experiments (Sherif & Sherif, 1953), there were significant intergroup issues that surfaced and had to be processed). The remainder of the agenda was devoted to movement through the various phases of the model described earlier using a combination of task and process initiatives. While the explicit goal of this session was the transformation of the individuals, groups and organization, it was also a parallel goal to avoid the shortfalls of many similar experiences that fail to "reintegrate" the process. This interest also continued in the next phase.

<u>The Sustaining Phase</u>. Conducted by each of four working groups that formed on the final day of the Applications Phase, this strategy was designed to overcome specific barriers to achieving the "envisioned organization." This process is continuing as of early 1984.

<u>Results</u>

The results of the application of this model can only be evaluated qualitatively, a concern to be addressed in the following section. In a report prepared for an independent evaluation, three key results were noted. First, the implementation plan

proposed by the working groups resulted in the establishment of three cross-divisional and multi-level committees. In many bureaucracies, the establishment of committees might be viewed as a step backwards, but in an organization characterized by a dearth of internal communication, this is a positive outcome. As one senior staff member said, "Before, we couldn't even have meetings. Now we can agree to disagree." Secondly, the spontaneously generated group of managers who formed to develop this process have been officially sanctioned by the commissioner to continue their efforts. Thirdly, the organization has formally adopted and integrated the model for change. Perhaps of more significance, the organization decided that the process was so valuable for their senior managers that they requested additional funding to provide the experience for all departmental managers. Adding independent support to the positive qualitative evaluation is the fact that this request for continuation was the only request fully funded in recent state budget action.

The Dilemma of Transformal Evaluation

The author entered the consulting relationship hoping to achieve three goals simultaneously. Naturally, a concern for the betterment of the client organization was paramount. Beyond that, it seemed to offer an opportunity to operationalize the model and, thirdly, it offered the chance to do research with an "external strategy" (Vroom, 1983). The first two goals were accomplished but the research objective seemed to become more and more entangled.

Anyone who has attempted action research is aware of the trade-offs in rigor that must sometimes be made. The consultants in this case were well aware of the "unintended consequences of rigorous research" (Argyris, 1970) and tried to balance "what we could learn" with "what we could do" whenever possible. Yet the difficulties of conducting research in an organization in need of transformal change seem far greater than one might expect in an organization only requiring transactional assistance. For example, the committee proposed that the Work Environment Scale (Moos, 1981) be used to provide a quantitative measure of effect. This scale was mailed to all participants with an explanatory letter from the consultants. But even under guaranteed anonymity, the response rate was so low that there could be no confidence attached to the results even though the direction was positive. As one regional non-respondent explained, "Even with that cover letter, I saw that survey and thought, 'here's more of that Central Office crap,' so I put it right in the trash can."

In a more generic sense, the dilemma involves the role of organizational behaviorists. Certainly, as the title of Professor Hackman's address states, we should be committed to "developing usable knowledge" (Hackman, 1984). Another responsibility involves assisting organizations in society. As a result of our opportunity and obligation to reflect on the nature of behavior in organizations, we may best be in positions to assist organizations as they attempt new and creative changes. These two roles

are not necessarily incompatible. However, when the needed process involves transformal change, as noted earlier, there is the possibility if not the likelihood that the organization is experiencing difficulty and this condition may reduce the likelihood of it offering itself up as an object for rigorous research. For the organization in this study, a qualitative analysis may have been the only methodology practical.

In a way, attempting to assist and study an organization in need of transformal change shares many of the problems faced by researchers as they searched for the causes and cures for malaria and yellow fever. The foul smelling, mosquito infested swamps were hardly the ideal location to conduct bacteriological research. Unfortunately, that was the only place the problem existed and was precisely the location where the treatment was needed. Similarly, organizations might not need transformal change if things are going fine. But in a troubled organization in need of transformation, while the environment for research might not be the best, it may well be the only one available. As Hackman has noted, "the theory and method required for research on organizational issues will have to be tailored to the special circumstances of the social systems in which the research is conducted---and to which the results are intended to apply. The theory and method of experimental or social psychology were not designed for this task, and they are insufficient for it." (in press).

Difficulties such as these should not be permitted to dissuade us from conducting this type of intervention or research based upon an external strategy. Vroom claims that, "it is organization problems and the desire of organizational members to do something about them that constitutes the most promising springboards for research" (1983, p.5). Clearly, the need for transformal change exists as does the requirement for an appropriate research methodology. Perhaps the field of organizational behavior could benefit from some transformation in its own research focus.

References

Argyris, C. Intervention theory and method. Reading, MA: Addison-Wesley, 1970.

Burns, J.M. Leadership. New York: Harper & Row, 1978.

Hackman, J.R. Psychological contributions to organizational productivity: A commentary. In A. Brief (Ed.), Research on productivity, New York: Praeger, In press.

Hackman, J.R. Developing usable knowledge about team effectiveness: Requirements for theory and method. Paper presented at 9th Psychology in the DoD Symposium, Colorado Springs, April, 1984.

Kouzes, J.M. & Mico, P.R. Domain theory: An introduction to organizational behavior in human service organizations. Journal of Applied Behavioral Science, 1979, 15(4), 449-469.

Lewin, K. Field theory in social science. New York: Harper & Row, 1951.

Mintzberz, H. The manager's job: Folklore and fact. Harvard Business Review, July-August 1975, 49-61.

Moos, R.M. Work Environment Scale. Palo Alto: Consulting Psychologists Press, 1981.

Sherif, M. & Sherif, C.W. Groups in harmony and tension. New York: Harper & Row, 1953.

Zaleznik, A. Managers and leaders: Are they different? Harvard Business Review, May-June 1977, 67-78.

COMPARATIVE EFFECTS OF BREMSSTRAHLUNG, GAMMA, AND ELECTRON RADIATION ON RAT MOTOR PERFORMANCE

Victor Bogo
Behavioral Sciences Department
Armed Forces Radiobiology Research Institute
Bethesda, Maryland 20814

ABSTRACT

The effects of rapidly delivered supralethal doses of bremsstrahlung, electron, and gamma radiation were investigated on the performance of male Sprague Dawley rats exposed at a midline tissue dose (MTD) rate of 2000 rad/min. The bremsstrahlung and electron radiations were produced in a linear accelerator (LINAC), and the gamma radiation was produced in a ^{60}CO facility. Relative radiation effects were determined by establishing median effective doses (ED50) for rats trained on the accelerod, a shock-avoidance test of motor performance. ED50's were based on 10-min post-exposure performance. The ED50 was 8121 rad for the bremsstrahlung field, 6110 rad for the electron radiation, and 3896 rad for the gamma photons. No significant difference was found between the bremsstrahlung and gamma radiations even though the physical characteristics used to achieve the two fields were less uniform in the bremsstrahlung exposures. However, the other ED50's translate into significantly different relative effectiveness of 1.35 between the bremsstrahlung and electron fields, and 1.45 between the electron and gamma radiations. The data imply that different radiation fields are not equally effective on behavior.

INTRODUCTION

High, rapidly delivered doses of ionizing radiation produce an immediate, transient decrement in performance (PD). All qualities of ionizing radiation, however are not equally effective at producing PD. Several investigators have shown that gamma photons are more effective at producing PD than are neutrons (George, 1971; Young, 1979). Other research has indicated that neither gamma photons, nor bremsstrahlung radiation is as effective at producing PD as is electron radiation (Bogo, 1983; Hunt, 1983).

PD has been demonstrated in a number of animal models performing different tasks after exposure to various radiation qualities (Bogo, 1983; George et al., 1971; Young, 1979). However, no comparison of PD has been made in one animal model performing the same task. These comparisons would address DoD-relevant issues about the relative biological effectiveness (RBE) of different radiation qualities, which would be useful in calculating personnel casualty criteria for functional integrity. Previously, the effects of bremsstrahlung and electron radiation on performance were compared (Bogo, 1983). The present study includes the gamma photon comparison. The objectives were (a) to

establish the median effective dose (ED50) for rat motor performance for each radiation (ED50 is defined as the dose at which 50% of the sample demonstrated PD), and (b) to establish the RBE for the three radiations tested (RBE defined as the ratio of the absorbed dose from one radiation to that of a reference radiation required to produce an identical biological effect under the same conditions) (Young, 1979).

METHODS

SUBJECTS. Subjects were 145 male Sprague Dawley rats that weighed 433g ± 4. They were individually housed in polycarbonate cages and maintained in keeping with the principles enunciated in the Guide for the Care and Use of Laboratory Animals prepared by the Institute of Laboratory Animal Resources, National Research Council. Food and water were available ad lib.

TASK AND TEST PROCEDURE. Rats were trained to maintain position on an accelerating rod elevated above a grid-shock floor, for as long as possible (Bogo et al., 1981). The accelerod rotational speed increased at an average of 0.9 rpm/sec. Subjects avoided shock by staying on the rod for >30 sec. The task was scored as duration of performance in seconds. Sham radiation conditioning lasted for 4 days, simulating all aspects of the radiation-day profile. Following irradiation, each subject was tested at 10, 15, 30, 60, and 120 min from the start of exposure and, when possible, at 24 hr. This paper deals only with performance on the 10-min test interval.

RADIATIONS. Rats were irradiated in a LINAC by either 18.6-MeV electrons (n = 40) or bremsstrahlung (n = 57) produced by 18.1-MeV electrons, or in a ^{60}CO facility by 1.25-MeV gamma photons (n = 48). Rats were individually irradiated, right side to the source, at a MTD rate of 2000 rad/min. LINAC dosimetry is discussed in greater detail by Bogo (1983).

ANALYSIS. The criterion measure was PD during the initial 10-min test period after exposure, which was determined using standard z scores. Z scores for each subject was obtained by dividing the difference between a postscore and a mean baseline score by the standard deviation of the animal's prescore values. PD was defined as performance at 2 z-scores below baseline. The initial doses for the ED50 estimate was determined for all radiation fields with an up/down sensitivity procedure (Dixon and Massey, 1969). An increased number of subjects were tested above and below the estimate to improve the confidence limits of the dose-response relationship. Probit analysis was applied to the obtained doses to determine the ED50 (Finney, 1978).

RESULTS

The number of subjects demonstrating PD on the 10-min test period by radiation at each dose level is presented in Table 1. While the dose levels to produce PD differed between radiation qualities, Table 1 shows that in all cases a dose-response relationship existed.

TABLE 1

SUBJECTS SHOWING PD FOR EACH RADIATION QUALITY

BREMSSTRAHLUNG			ELECTRON			GAMMA		
DOSE	PD/N*	%PD	DOSE	PD/N	%PD	DOSE	PD/N	%PD
6400	1/6	17	4600	0/8	0	5900	1/12	8
7700	5/13	38	5200	2/8	25	8700	5/14	36
8700	14/22	66	5900	2/8	25	9700	8/10	80
9900	15/16	94	6700	6/8	75	13000	9/10	90
			7700	8/8	100	17000	2/2	100

* N = Number of subjects in group

Probit analysis was made assessing the effects of each radiation quality. Since no significant differences were found in the three slopes, a pooled estimate was used to construct the probits shown in Figure 1. The probits indicate that the electron field affected performance at lower doses than did the other fields, and that bremsstrahlung and gamma were almost equivalent. The ED50 was 8121 rad for the bremsstrahlung field, 6110 rad for the electrons, and 8896 rad for the gamma photons. No significant difference was found between the bremsstrahlung and gamma fields: however, the ED50's for the other comparisons translate into significantly different RBE's of 1.35 between the bremsstrahlung and electron radiations, and 1.45 between the electrons and gamma photons.

DISCUSSION

The present study is the first to compare the effects of bremsstrahlung, electron and gamma radiation on the behavior of one animal model performing the same task. Electron radiation disrupted performance at significantly lower doses than did the bremsstrahlung and gamma fields. These findings agree with those reported by Hunt (1983), who found that electrons affected rat performance at lower doses than did gamma photons. In addition, the present finding of no difference between the bremsstahlung and gamma fields also agrees with the literature reporting that these fields are equivalent (Young,1981).

Other studies have reported different effects of radiation on performance; however, the reason for the difference is not readily apparent. In these studies, bremsstrahlung or gamma radiation was found to be more effective than neutron radiation (George et al., 1971; Young, 1981). Young (1981) speculated that the difference may be related to low-LET photons having a greater effect on the available oxygen; thus, the oxygen-rich CNS would be more affected. However, no such difference existed in the present study; therefore, this possible explanation does not shed any light on the differences.

Another explanation for the present findings are differences in the physical characteristics used to achieve the radiations.

As shown in the Methods section, the energy levels were unequal; however, the LET's for the fields are almost the same (ICRU, 1970). Thus, the energy level differences do not explain the behavioral findings. In addition, the bremsstrahlung exposures were not as uniform as were the electron and gamma fields in terms of depth dose and dose delivered to the head. However, at this time, it is not known how much to attribute to this factor. Since the present study confirms previous research (Hunt, 1983) and since bremsstrahlung and gamma radiation are supposed to be equivalent, it is questionable that this uniformity difference will have a marked bearing on the present findings. The main consideration however is that even with this speculation, the present research indicates that different radiation fields are not equally effective at disrupting behavior.

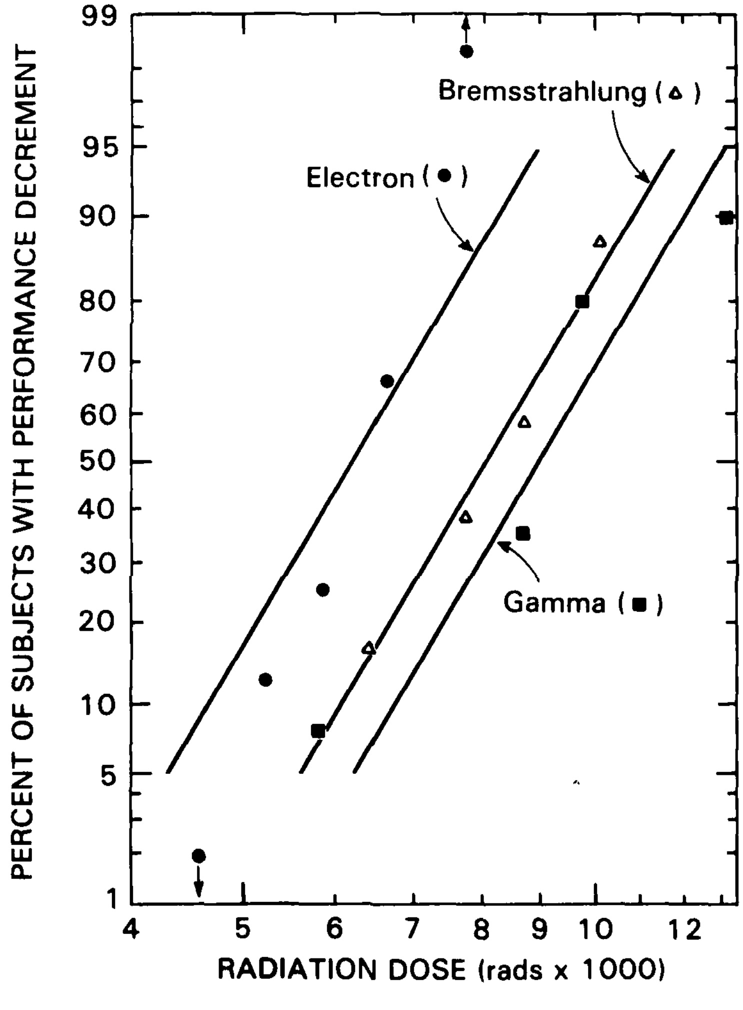


Figure 1. Accelerod performance decrement dose-response curves produced by bremsstrahlung, electron, and gamma radiation. Circles, squares and triangles represent the mean data point for the respective radiation quality.

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of C. Boward, W. Jackson, G. Kessell, and AFRRI's dosimetry division. This research was supported by the Armed Forces Radiobiology Research Institute, Defense Nuclear Agency (DNA), under work unit MJ 00096. Views presented are those of the author; no endorsement by DNA has been given or should be inferred.

REFERENCES

Bogo, V. Effects of bremsstrahlung and electron radiation on rat motor performance; In press, 1984.

Bogo, V., Hill T.A., and Young, R.W. Comparison of accelerod and rotarod sensitivity in detecting ethanol- and acrylamide-induced performance decrement in rats: Review of experimental considerations of rotating rod systems. Neurotoxicology 2, 765-787 (1981).

Dixon, W.J. and Massey, F.W. Jr.. Introduction to Statistical Analysis. McGraw-Hill, New York (1969).

Finney, D.J. Quantal responses and the tolerance distribution. In Statistical Methods in Biological Assay, pp. 349-369. McMillan Publishing Co., Inc., New York, 1978.

George, R.E., Chaput, R.L., Verrelli, D.M. and Barron, E.L. The relative effectiveness of fission neutrons for miniature pig performance decrement. Radiation Research 48, 332-345 (1971).

Hunt, W.A. Comparative effects of exposure to high-energy electrons and gamma radiation on active avoidance behaviour. International Journal of Radiation Biology 44, 257-260, 1983.

International Commission on Radiation Units and Measurements. Linear Energy Transfer. Report No. 16, ICRU Publications, Washington, DC, 1970.

Thorpe, J.W. and Young, R.W. Monkey performance after partial body irradiation. Aerospace Medicine 42, 503-507 (1971).

Young, R.W. Prediction of the relative toxicity of environmental toxins as a function of behavioral and non-behavioral endpoints. Dissertation submitted to The Catholic University of America, Washington D.C., 1979.

AD-P003 251

The Interactive Effects of Morphine and Ionizing Radiation on the Latency of Tail-Withdrawal from Warm Water in the Rat

Walter F. Burghardt and Walter A. Hunt
Armed Forces Radiobiology Research Institute
Behavioral Sciences Department

Abstract

The analgesic effect of morphine (1-4 mg/kg, i.m.) was enhanced in rats exposed to ionizing radiation (250-5000 rad) in a dose-dependent manner. This was indicated by an increase in the latency of tail-withdrawal from warm (56 °C) water, compared with animals receiving morphine alone. Radiation alone had no effect on latencies or on gross behavior. The enhancing effect of radiation was strongest at 24 hours after irradiation and was partially reversible with naloxone, an opiate antagonist, at a dose of 2 mg/kg, i.m. No changes in survival time after irradiation were noted between animals receiving morphine and those receiving saline injections. The results of this study suggest that the effect of narcotic analgesics to relieve pain in casualties on a nuclear battlefield may be enhanced depending on the postirradiation interval.

Introduction

The actions of a number of classes of pharmaceutical agents appear to be altered after exposure to ionizing radiation (Barnes, 1967; Davis et al., 1968; Nair, 1967). The effects of ionizing radiation on the actions of the narcotic analgesics, however is somewhat unclear, due to the paucity of research in the area, and the conflicting results of some published studies (eg. Kocmierska-Grodzka and Kozlowska, 1972; Doull, 1967; Mickley et al., 1983). The reasons for the difference in results are not clear, but there were a number of differences between the experimental designs, including task used, time of testing, dosage and route of administration of the agents, and dose and qualities of radiation used.

The problem of measuring the efficacy of the narcotic analgesics requires more work over a range of doses and an attempt to control for factors which differed significantly among the earlier studies. The current study incorporated a non-invasive, accepted, reliable, and relatively non-damaging method for the determination of the analgesic capabilities of morphine (Miksic and Lal, 1977) at various times after irradiation, using a number of doses of morphine and of mixed neutron/gamma radiation.

Methods

The subjects were male Sprague-Dawley rats (300 g). The animals were experimentally naive, and were fed and watered *ad lib*. The animals were kept on a reverse 12-hour light cycle.

The apparatus consisted of a clear plexiglas restraining tube which allowed the rat's tail to hang loosely when the animal was in the tube. The water bath was maintained at $56 \pm 1^{\circ}C$, using a laboratory hot plate. A stand was placed next to the water bath on which the restraining tube and animal were placed during experimental trials. This allowed for controlled, 3-4 cm insertion of the tail in the water. Latency of tail-withdrawal was measured using a contact relay, electromechanical timer, and two electrodes (one placed in the water bath and the other fastened to the base of the animal's tail using electrode paste and adhesive tape rings). The animal's tail was removed from the water if no response was made in 15 seconds, and maximum latency was, therefore, defined as 15 seconds.

On the day before treatment, the animals were placed in the restraining tube, and allowed 15 minutes to accommodate. Ten trials of tail-withdrawal were run with 2-minute intertrial intervals. This session served as their baseline. The animals were then returned to their home cages. On the experimental day, the procedure was the same, except as noted below. Three separate experiments were conducted.

In the first study, experimental animals received 5000 rads of mixed neutron/gamma radiation (neutron:gamma, 3:1) 24 hours after baseline determination. Control subjects were not irradiated, but were otherwise treated the same. Six irradiated animals each were randomly assigned to one of 8 experimental groups. Four of the irradiated groups received 4 mg/kg morphine intramuscularly at 1, 2, 4, or 24 hours after irradiation, the other four groups received saline injections at the same times. Thirty minutes after the injection, the experimental groups were tested for 10 trials of tail- withdrawal. Two control groups of 6 animals each were retested 24 hours after baseline determination and 30 minutes after saline or 4 mg/kg morphine injections, i.m. In experiment 2, the dose of irradiation was decreased to 500 rads. The technique was similar to the first study, except that there were only three experimental groups of 6 animals each, injected with 4 mg/kg morphine, i.m., at 1, 24, or 48 hours after irradiation.

The third study was similar to the the above two studies. In this experiment, there were 3 groups of 6 animals. All subjects received 500 rads, and all were injected at 24 hours after irradiation. The groups differed in the dose of

morphine received (1, 2, or 4 mg/kg, i.m.). The animals were given 10 trials of tail-withdrawal testing, 30 minutes after injection. All subjects were then given 2 mg/kg naloxone, i.m., and retested for 10 trials, 10 minutes after injection.

Data below are expressed as peak tail-withdrawal latency for each treatment. Data were analyzed using a one factor analysis of variance (Winer, 1962).

Results

The first study showed that 5 krad alone did not produce a significant change in tail-withdrawal latencies at 1, 2, 4, or 24 hours after irradiation when compared with nonirradiated controls ($p>.1$). Morphine alone at 4 mg/kg produced a significant increase in tail-withdrawal time (244% baseline). Peak tail-withdrawal latencies of irradiated animals receiving 4 mg/kg morphine were not significantly different from animals receiving morphine alone at 1, 2,or 4 hours after irradiation, although the group injected at 24 hours showed a significant difference from the unirradiated morphine group (564% baseline). Differences cited above were significant at $p<.01$ ($F_{9,50}=17.143$). There were no significant differences between the survival times of any of the irradiated groups.

The second study demonstrated that at 500 rads, 4 mg/kg morphine at 1 hour and 48 hours did not produce a significant increase in tail-withdrawal latency compared with control animals. However, at 24 hours, morphine produced an increase in tail-withdrawal which was significantly different from all other groups (472% baseline). Differences were significant at $p<.01$ ($F_{3,20}=42.689$).

The final study indicated that 24 hours after 500 rads, morphine produced increases in tail-withdrawal latency in a dose dependent manner. Four mg/kg increased latency to 336 percent of baseline, 2 mg/kg increased latency to 250 percent, and 1 mg/kg increased latency to 37 percent. Naloxone at 2 mg/kg, i.m., returned the response to 4 mg/kg to 167 percent of baseline, and completely reversed the effects of the lower doses of morphine ($p<.01$, $F_{6,35}=3.538$).

Discussion

Although the results of the above studies are only preliminary, several conclusions may be tentatively drawn. First, at about 24 hours after irradiation as low as 500 rads, there appears to be an increased sensitivity to the analgesic effects of morphine. The increase is not evident

between 1 and 4 hours, and is no longer present at 48 hours. Irradiation alone produces no change in tail-withdrawal latency, and the difference between irradiated animals receiving morphine and those receiving saline is not reflected by a change in survival times. Finally, the analgesic effect appears to be at least partially reversible by naloxone.

If these results are borne out in further research, they would appear to reconcile some of the earlier studies which showed conflicting results, depending on when the narcotic was administered after radiation. In addition, results of this and subsequent studies may provide recommendations for the administration of narcotic analgesics in a combined injury victim on the battlefield.

References

Barnes, C.D. Central nervous system drugs and x-irradiation: their interactive effects. Radiation Research, 1967, 30, 351-358.

Davis, L.W., Brown, J.A., and Strike, T.A. Effectiveness of drugs in animals exposed to mixed gamma-neutron radiation: V. Barbiturates. AFRRI Report 69-19, 1968.

Doull, J. Pharmacological responses in irradiated animals. Radiation Research, 1967, 30, 333-341.

Kocmierska-Grodzka, D. and Kozlowska, M. Badania nad dzialaniem przeciw-bolowym morfiny w przebiegu choroby popromiennej myszy. Acta Physiologica Polonica, 1972, 23, 149-157.

Mickley, G.A., Stevens, K.E., Burrows, J.M, White, G.A., and Gibbs, G.L. Morphine tolerance offers protection from radiogenic performance defecits. Radiation Research, 1983, 93, 381-387.

Miksic, S. and Lal, H. Tolerance to morphine produced discriminative stimuli and analgesia. Psychopharmacology, 1977, 54, 217-221.

Nair, V. Modification of pharmacological activity following x-irradiation. Radiation Research, 1967, 30, 359-373.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Quaternary Naltrexone Reverses Radiogenic and Morphine-induced Locomotor Hyperactivity

G. Andrew Mickley
Armed Forces Radiobiology Research Institute
Bethesda, MD 20814

Karen E. Stevens
University of Colorado
Colorado Springs, CO 80933

James A. Galbraith
United States Air Force Academy
Colorado Springs, CO 80840

Gerald A. White and Gregory L. Gibbs
Penrose Cancer Hospital
Colorado Springs, CO 80903

Abstract

Previous research has indicated some similarities between behaviors observed after treatment with morphine and exposure to ionizing radiation. These data suggest that endogenous opiate substances (endorphins) may be involved in radiogenic behavioral change. The present study attempted to determine the relative role of the peripheral and central nervous system in the production of morphine-induced or radiation-induced locomotor hyperactivity of the C57BL/6J mouse. Toward this end, we used a quaternary derivative of an opiate antagonist (naltrexone methobromide), which presumably does not cross the blood-brain barrier. Quaternary naltrexone (40 mg/kg, i.p.) was used to challenge the stereotypic locomotor response observed in these mice after either an i.p. injection of morphine or exposure to 1500 rads ^{60}Co. The quaternary derivative of naltrexone reversed the locomotor hyperactivity normally observed in the C57BL/6J mouse after an injection of morphine. It also significantly attenuated radiation-induced locomotion. The data reported here support the hypothesis of endorphin involvement in radiation-induced locomotor hyperactivity and, taken at face value, they suggest an important role for peripheral opiates in morphine-induced and radiogenic behaviors. However, these conclusions are contingent upon further research which more fully evaluates naltrexone methobromide's capacity to cross the blood-brain barrier.

Introduction

When C57BL/6J mice are exposed to a sufficiently large dose of ionizing radiation, they exhibit a dose-dependent, stereotypic locomotor hyperactivity that is, in some ways, similar to that observed in morphine-injected mice (Mickley, Stevens, White and Gibbs, 1983a; Olivero and Castellano, 1973; Teitelbaum, Giammatteo and Mickley, 1979). This radiogenic behavior can be reversed by an injection of the opiate antagonist naloxone

(Mickley et al., 1983a). These data, along with others (Mickley, Sessions, Bogo and Chantry, 1983b; Mickley, Stevens, White and Gibbs, 1983c; Mickley, Stevens, Burrows, White and Gibbs, 1983d), suggest that the endogenous opiates (endorphins) may play some role in radiation-induced behavioral alterations.

The locus, as well as the exact mechanism(s) of action of these radiogenic endorphins, remain unclear. Therefore, the present studies attempted to specify the extent of peripheral or central nervous system involvement in the production of morphine- and radiation-induced locomotor hyperactivity. Toward this end, we used a quaternary derivative of naltrexone (an opiate antagonist), which presumably does not cross the blood-brain barrier (Brown,Robertson and Goldberg, 1983; Miczek, Thompson and Shuster; Ferretti, Bianchi, Tavani and Manara, 1981; Russell, Bass, Goldberg, Schuster and Merz, 1982). Specifically, we hypothesized that if radiogenic or opioid-induced locomotion are reversed by an injection of this peripherally acting antagonist, it can be inferred that these behaviors are being at least partially mediated by neurons outside the brain and spinal cord.

Methods

Subjects

The subjects were male C57BL/6J mice (15-20 g) obtained from Jackson Laboratories (Bar Harbor, Maine). Subjects were placed in individual cages at least 24 hours before activity testing was begun. Mice were maintained on a 12-hour light-dark cycle (lights on 6:00 A.M.) in a temperature-controlled room ($72^{\circ} \pm 3^{\circ}$). Purina laboratory chow and tap water were available ad libitum except during behavioral measurements and drug injections.

Procedures

Twenty mice received an injection (i.p.) of 30 mg/kg morphine sulfate while 20 other control animals were injected (i.p.) with an equal volume of saline. Morphine has been shown to produce a stereotypic locomotor hyperactivity in the C57BL/6J mouse (Olivero and Castellano, 1974; Teitelbaum et al., 1979). Five minutes after these injections locomotor activity was recorded during two successive 15-minute periods, using the Automex D system of Columbus Instruments. Half of the animals in the morphine and saline groups then received an injection of naltrexone methobromide (40 mg/kg, i.p. in 0.9% saline). This dose of quaternary naltrexone was chosen since it has been shown that this opiate antagonist is 40 times less potent than its tertiary counterpart (Brown et al., 1983). The remaining subjects were injected with an equal volume of saline. After a 5-minute wait, locomotor activity was recorded for an additional two 15-minute periods.

In a second experiment we measured locomotor activity in 40

irradiated or sham-irradiated mice. Initially, two 15-minute periods of pre-irradiation/sham locomotor activity were recorded 5 minutes after a control saline injection. The animals were then confined, and half were exposed to 1500 rads ^{60}Co (see procedure in Mickley et al. 1983a). This dose of radiation has been shown to produce a locomotor hyperactivity in the C57BL/6J mouse (Mickley et al. 1983a; Mickley et al. 1983b). The remaining subjects were sham-irradiated. Immediately after irradiation, half (N = 10) of the mice received an injection of 40 mg/kg naltrexone methobromide. The remaining irradiated mice (N = 10) received a control saline injection. Sham-irradiated subjects were also treated with 40 mg/kg naltrexone methobromide (N = 10) or saline (N = 10). Beginning 5 minutes after irradiation/sham, the locomotor activity of these subjects was again recorded for two 15-minute periods.

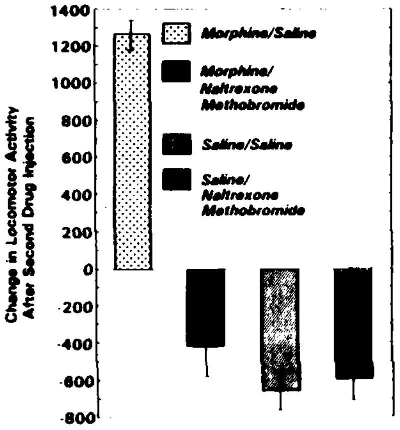
Results

Quaternary naltrexone reversed the locomotor hyperactivity normally observed in the C57BL/6J mouse after an injection of morphine. As has been reported in previous studies (Teitelbaum et al., 1979), 30 mg/kg morphine sulfate produced a significant locomotor hyperactivity as compared to controls receiving only saline, $t(18) = 8.05$, $p < 0.001$. In a further analysis we computed, for each animal, a "change statistic," which reflected the total locomotor activity observed over the 30 minutes after the second injection (quaternary naltrexone or saline) minus the baseline activity in the same period of time after the initial injection of morphine or saline (see Figure 1). These numbers were used in the calculation of a t test (Winer, 1971) comparing the locomotor activity of morphine-injected mice with that of opiate-treated mice that had also received quaternary naltrexone. Although the initial morphine-induced locomotion (in the first 30 minutes) was similar in these two groups, treatment with naltrexone methobromide reduced activity to levels significantly lower than those observed in saline-injected controls, $t\ (18) = 4.67$, $p < 0.001$. Quaternary naltrexone alone apparently does not produce hypoactivity, since a comparison of locomotor activity counts between saline-injected subjects and those receiving naltrexone methobromide revealed no statistically significant difference between these groups, $t\ (18) = 0.07$, $p > 0.05$.

Quaternary naltrexone also significantly attenuated radiation-induced locomotor hyperactivity. Pre-irradiation/sham baseline activity measures were not significantly different between the groups in this study [one-way ANOVA; $F\ (3,36) = .571$, $p > 0.05$] (Winer, 1971). Therefore, to analyze the radiation data, we computed a "change statistic," which reflected the total locomotor activity during the first 0.5 hour after irradiation or sham exposure minus the total activity recorded in the first 0.5 hour of pre-irradiation/sham baseline. An initial one-way analysis of variance revealed the existence of significant radiation-dependent or drug-dependent differences between these changes in mouse locomotion, $F\ (3,36) = 2.894$, $p = .05$. Planned

individual comparisons suggested that saline-injected subjects increased their activity to a greater extent after irradiation than they did after sham irradiation, t (36) = 2.089, p <.04. As Figure 2 illustrates, this radiogenic locomotor hyperactivity was significantly reduced by an injection of naltrexone methobromide, t (36) = 2.735, p =.01. In fact, the activity of irradiated mice injected with the peripheral opiate antagonist was not significantly different from the locomotor response of those subjects that were sham-irradiated and had received injections of physiological saline, t (36) = 0.645, p >0.05. Apparently this reduction in locomotion does not reflect a generalized lethargy caused by quaternary naltrexone. Sham-irradiated mice that received the drug were not statistically distinguishable from those that received a saline control injection, t (36) = 0.101, p >0.05.

Fig. 1



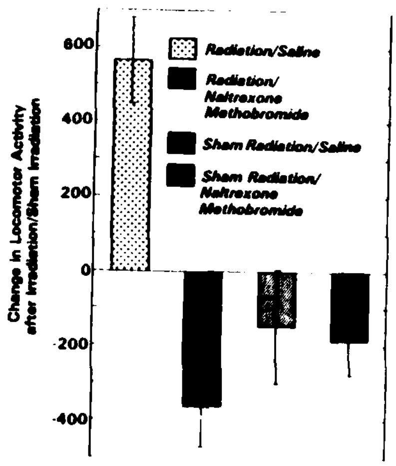


Fig. 2

Figure 1. Changes in locomotor activity after an initial injection of either morphine (30 mg/kg) or saline followed by a second injection of either quaternary naltrexone (40 mg/kg) or saline. Naltrexone methobromide reversed morphine-induced locomotor hyperactivity to the extent that it was indistinguishable from the behavior of control mice.

Figure 2. Locomotor changes (from pre-irradiation baseline evels) in irradiated and sham-irradiated mice injected with quaternary-naltrexone or saline. Naltrexone methobromide reversed radiogenic hyperactivity.

Discussion

Naltrexone methobromide effectively challenged morphine-induced and radiogenic locomotor hyperactivity. The extent of this reversal of hyperactivity seemed to be comparable to that already reported after an injection of naloxone (Mickley et al., 1983a). An examination of these data seems to support a peripheral site of action mediating these opioid-induced behaviors. However, this interpretation should not be made without some qualification. First, a strong data base already exists, suggesting that morphine-induced locomotion is mediated, at least in part, by the central nervous system (Brown,

Derrington and Segal, 1979; Olivero, 1974; Van Abeelen and Van den Heuvel, 1982; Rogers and File, 1979; Teitelbaum et al., 1979). Second, despite a growing number of studies utilizing naltrexone methobromide to determine the sites at which opiates exert their effects, this drug has yet to be fully evaluated for its ability to cross the blood-brain barrier. (Bianchi, Fiocchi, Tavani and Manara, 1982) (Brown et al., 1983). It may be the case that, over time, quaternary naltrexone may leak into the brain or alternatively it may be metabolized into naltrexone. Perhaps it is naltrexone (which crosses the blood-brain barrier more readily) that antagonizes morphine's and radiation's effects (Brown et al., 1983). The results presented here do not resolve these issues.

Although radiation exposure is known to alter the blood-brain barrier, the dose and latency required to observe these changes in post-irradiation permeability are unclear (Clemente and Richardson, 1962). To the best of our knowledge no radiogenic alterations in the blood-brain barrier have been reported within the limited post-irradiation time period used in the present experiments (5-35 minutes post exposure). This, of course, does not confirm that changes in permeability did not occur. It is interesting to note here, however, that quaternary naltrexone reversed morphine-induced locomotion (when the blood-brain barrier was intact). This suggests that an intact blood-brain barrier may not eliminate naltrexone methobromide's antagonism of this opiate-induced behavior.

The data reported here support the hypothesis of endorphin involvement in radiation-induced locomotor hyperactivity and, taken at face value, they suggest an important role for peripheral opiates in morphine-induced and radiogenic behaviors. However, these conclusions are contingent upon further research which more fully evaluates naltrexone methobromide's capacity to cross the blood-brain barrier of the C57BL/6J mouse.

(Contact Author for References)

Acknowledgements

The authors wish to thank SSgt. Robert L. Stewart and Sgt. Joseph Dannelley for their fine technical assistance during the course of these experiments. We wish to thank Dr. H. Merz for his contribution of naltrexone methobromide. Penrose Cancer Hospital's superior cooperation in the radiation experiments is gratefully acknowledged. In this regard, Ms. Valerie Radice provided us with excellent technical assistance in the use of the ^{60}Co source. We thank Mrs. Marion Golightly for her help in typing this manuscript. Research was conducted according to the principles enunciated in the "Guide for the Care and Use of Laboratory Animals" prepared by the Institute of Laboratory Animal Resources, National Research Council. This experiment was conducted while the first author was at the USAF Academy and was supported by Defense Nuclear Agency (DNA) under research project U99QMXMK, 00041.

SYMPOSIUM LUNCHEON ADDRESS

A PERSONALITY PROFILE FOR GENERAL OFFICERS

Dr. David P. Campbell
(Center for Creative Leadership)

Dr. David Campbell is the Smith Richardson Senior Fellow in Creative Leadership at the Center for Creative Leadership, Greensboro, North Carolina. He is also an Adjunct Professor at Duke University.

The Center for Creative Leadership is a nonprofit institution established by the Smith Richardson Foundation to develop the knowledge of the behavioral sciences into useful applications for leaders and aspiring leaders. The Center carries on an extensive research program on leadership with emphasis upon the creative, non-routine elements.

Dr. Campbell, a native Iowan, received his B.S. and M.S. from Iowa State University and his PhD -- 1960 -- from the University of Minnesota. He was appointed to the Minnesota faculty in 1960 and became a full professor in 1968. In 1973 he was invited to the Center for Creative Leadership as a Visiting Fellow; six months later he accepted the newly-created position of Vice-President of the Center.

Dr. Campbell's earlier work centered around psychological testing and career development. He is the co-author of the Strong-Campbell Interest Inventory, the premier instrument of its kind. In more recent years, he has also been involved in research in leadership and creativity. Dr. Campbell has published widely in scientific and professional journals, and he has authored several best-selling books (the most recent being *If I'm in Charge Here, Why is Everybody Laughing?*, published in 1980). In addition to academic honors that include the E. K. Strong Gold Medal, he values the University of Minnesota Squash championship, and runner-up in the Minneapolis *Tribune* Annual Photo Contest. He is on the Board of Directors of National Computer Systems. His professional affiliations include membership and fellowship in the Division of Industrial and Organizational Psychology, the Division on Evaluation and Measurement, and the Division of Counseling Psychology of the American Psychological Association.

His luncheon address at the Ninth Biennial Psychology in the DOD Symposium on 18 April 1984 is titled "A Personality Profile for General Officers." A summary of Dr. Campbell's address follows.

A PERSONALITY PROFILE FOR GENERAL OFFICERS

David P. Campbell

Center for Creative Leadership

During the years of 1978 through 1981, roughly 1000 people went through the Center for Creative Leadership's Leadership Development Program in Greensboro, NC. Of these, 64 were Brigadier Generals in the U.S. Army. Each participant filled in a lengthy battery of psychological tests, and was also observed and rated on their performance in small group exercises.

This paper is a summary of the test scores and ratings for these military officers versus the other subgroups, and the total sample. The results, generally predictable, showed the military officers scoring high on scales such as Dominance and Willingness to take control. Of course, as in every sample, there was a lot of variability within the sample and perhaps the most interesting finding was the variations in the individual patterns in the test scores of these outstandingly successful officers.

PANEL SESSION

PANEL SESSION: RADIATION-INDUCED COMBAT PERFORMANCE DEGRADATION: THE DEFENSE NUCLEAR AGENCY INTERMEDIATE DOSE PROGRAM

SESSION CHAIR: G. Rufus Sessions (U.S. Air Force Academy)

PANELISTS: Robert W. Young (Defense Nuclear Agency)

Albert S. Glickman (Organization Research Group of Tidewater, Inc.)

George H. Anno (Pacific-Sierra Research Corporation)

G. Rufus Sessions (U.S. Air Force Academy)

PROCEEDINGS ENTRIES

"The defense nuclear agency intermediate dose program: An overview" (Robert W. Young and David L. Auton)

"A questionnaire assessment of estimated radiation effects upon military task performance" (Albert S. Glickman, Peter S. Winne, Ben B. Morgan, Jr., and Richard B. Moe)

"Estimated effects of ionizing radiation upon military task performance: Individual combat crewmember assessments" (George H. Anno and Don B. Wilson)

"A summary of the psychological effects of tactical nuclear warfare" (G. Rufus Sessions)

THE DEFENSE NUCLEAR AGENCY INTERMEDIATE DOSE PROGRAM: AN OVERVIEW (EFFECTS OF TOTAL-BODY IRRADIATION ON THE PERFORMANCE OF PERSONNEL IN ARMY COMBAT CREWS)

ROBERT W. YOUNG, PhD and DAVID L. AUTON, PhD
Defense Nuclear Agency

ABSTRACT

The objective of this research was to provide the quantitative basis for predicting performance decrement in Army crewmen irradiated with less than 4500 rads (cGy). Since the data necessary for these predictions do not exist and cannot be obtained directly by human testing due to the deadly nature of ionizing radiations, this effort involved not only the collection of data, but the development of a methodology to indirectly obtain the data. The data were obtained using a questionnaire derived from detailed information on radiation sickness and analysis of 15 combat crew tasks. The questionnaire, which asked for quantitative information on the impact of radiation sickness symptoms on the performance of sub-tasks, was administered to experts in the performance of the combat tasks. The results obtained in this effort clearly demonstrate that this methodology can be used to obtain meaningful estimates of the impact of very hazardous environments on performance. Comparison of the results from this study with those from studies which have directly assessed the effects of sickness on performance suggests that this questionnaire approach could successfully be applied to the evaluation of other hazardous environments in other military systems.

INTRODUCTION

The objective of this research effort was to provide the quantitative basis for predicting the performance decrement in Army infantry crewmen irradiated with less than 4500 rads (cGy). Army analysis of critical factors on the tactical battlefield indicates that nuclear radiation is the predominate casualty producing effect for nuclear weapons of 50 Kt or less. Present senerios for combat operations in a nuclear environment indicate that between one-half and three-quarters of the infantry personnel targeted with a tactical nuclear weapon would receive an initial radiation dose of greater than 150 rads (cGY) but less than 3000 rads (CGy). While, in general, the effects of radiation in this dose range are known, the precise effects of radiation on combat performance are not well known, especially the effects of radiation sickness. The present combat casualty criteria are based on the incapacitating and

killing effects of radiation derived primarily from animal experimental studies. Since no data exists with which to quantify degradation in performance from radiation sickness, these effects are not presently defined in the combat casualty criteria. While it is generally known that most of the individuals who receive more than the "Emergency Risk Dose" of 150 rads (cGy) but less than a "Latent Lethal Dose" of 650 rads (cGy), will experience radiation sickness, it is not known how radiation sickness will affect combat capability, despite the probable survival of many of these individuals. Since there is no way to directly test the performance of combat personnel in an actual or simulated radiation environment, obtaining data for these predictions is a major obstacle to accurate tactical nuclear planning.

The direct prediction of combat capability after a nuclear burst is not possible since there is no history of combat operations in a nuclear environment. In fact, there are very limited sources of information on the effects of prompt total-body irradiation--namely Hiroshima/Nagasaki, accidental exposures, clinical irradiations, and animal experiments. In none of these cases except the animal experiments was task performace measured after irradiation. There are no performance measures which can be reconstructed from Hiroshima and Nagasaki. Additionally, uncertainties about the time at which radiation sickness occurred and the dose which individuals received limits the use of information which can be obtained from those bombings. Likewise, none of the accidental exposures produce any estimate of the effects of radiation on performance, except for the one case in which early transient incapacitation was documented in the postirradiation clinical description (Shipman, et al., 1961). Only two studies (Payne, et al., 1963; Wolfgang and Maier, 1972) which have conducted as follow-ups to clinical irradiations have measured any aspect of psychomotor performance capability. Both of these studies employed severely ill cancer patients and neither employed performance measures which could be directly related to military task performance. Postirradiation performance has been measured in numerous animal experiments. This source of data provides good quantification for postirradiation incapacitation but much less information on the impact on performance of radiation sickness. Given the limitations on existing sources of data and the unacceptability of human testing to directly measure these effects, an indirect method had to be established to obtain measures of intermediate doses of ionizing radiation on combat performance.

APPROACH

A combination of several methods was employed in this project. The basic approach was to administer a questionnaire to senior crewmen in Army combat crews in order to obtain quantifiable judgements on the effects of radiation sickness upon the performance of their tasks. The questionnaire employed behaviorally anchored rating scales to obtain estimates of the impact of the symptoms of radiation sickness on the performance of

sub-task elements of Army combat crew functions. Prerequisite to the development and administration of the questionnaire was a detailed description of both radiation sickness Army crew functions. Subsequent to questionnaire administration were the multivariate analysis of the data, efforts to corroborate the results, and the development of a method to qunatitavively generalize to the data to the entire time-dose matrix. The discrete major elements which comprised this project were: (1) generating quantitative descriptions of radiation sickness, (2) developing behaviorally anchored scales of radiation sickness, (3) combining radiation sickness scales to produce "typical" symptoms complexes of illness syndromes, (4) combining radiation sickness syndromes and task analyses into a questionnaire, (5) administering the questionnaire, (6) applying a multivarate analysis to the results, (7) developing a quantitative method to generalyze the results to the entire time-dose matrix, (8) accounting for the psychological impact of nuclear combat, (9) integrating animal incapacitation data into the human performance decrement data, and (10) applying the analysis to Army planning. As such, this study was a multidisciplinary effort accomplished by several different organizations and guided by a working group.

The quantification of the signs and symptoms of radiation sickness during the first 6 weeks after exposure was undertaken as the first step in this project. To facilitate this process, the range of radiation doses between 75 and 4500 rads (cGy), free-in-air, was subdivided into 8 dose ranges associated with discrete pathophysiological events. For each subrange of dose, the "typical" incidence, severity, and duration of radiation sickness signs and symptoms were defined. These endpoints were quantified from the medical and research literature on human radiation sickness. This literature consists of (1) case studies of nuclear accidents, (2) records of patients given radiotherapy for cancer and other diseases, (3) analysis of "composite" studies, including the experience of the Japanese atomic bomb survivors, and (4) expert opinion. These data were analyzed by a working group of experts in radiobiology and radiation therapy to reach a consesus about "typical" symptoms.

In order to incorporate this radiation symptom information into behaviorally anchored rating scales for use in a questionnaire, the radiation symptoms information was divided into 6 separate categories of acute radiation symptoms. The symptom categories were: (1) upper gastrointestinal destress, (2) lower gastrointestinal distress, (3) fatiguability and weakness, (4) hypotension, (5) infection and bleeding, and (6) fluid loss and electrolyte imbalence. The symptoms for each of the categories of radiation-induced illness were described for mild to severe symptoms on a 5-point rating scale. Symptom complexes were then constructed, based on the severity descriptions for the 6 symptom categories and their occurance within the 8 radiation dose ranges. Since radiation sickness is manifest in a time-dependent manner, illness syndromes were simply a combination of symptom descriptions for each symptom category that occurred concurrently in time. From

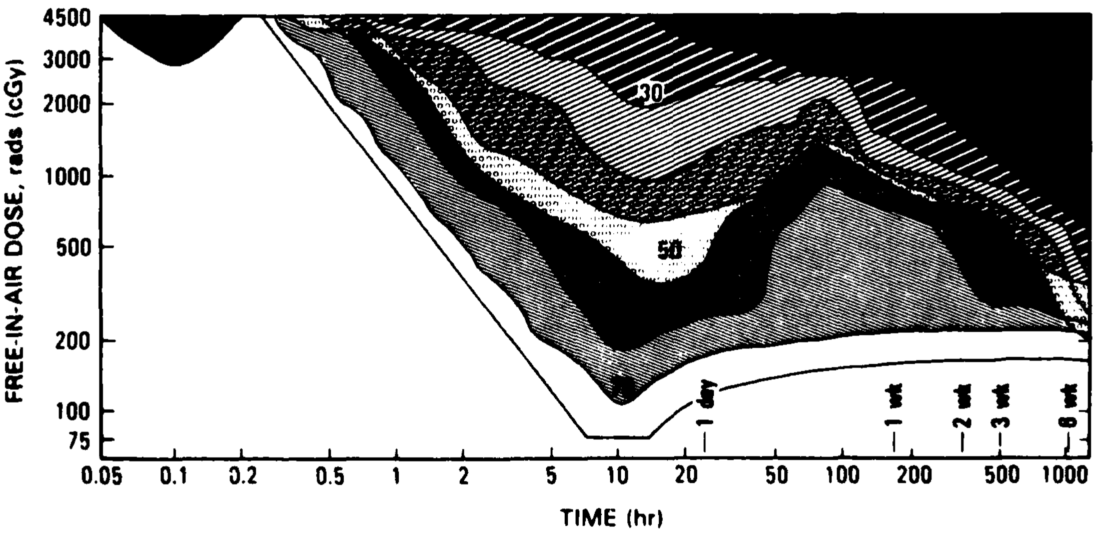
the large number of possible symptom complexes, the 30-40 unique syndromes which provided the greatest coverage of the time-dose matrix were chosen. As such, syndromes ranged from mild symptoms from a single category to to severe symptoms from several categories.

Each item on the questionnaire was comprised of a radiation sickness syndrome and a task element from the crew functions being tested. Respondants were asked to estimate the time required to complete a subtask element under the conditions of illness specified in the question. The impact of radiation sickness on each crew position was assessed in this way for the M60A3 tank, 155 mm self-propelled howitzer and fire-direction center, and the improved TOW vehicle. A detailed description of the construction and administration of the questionnaire is reported by Winne, et al. (1984) elsewhere in this symposium. Likewise, the analysis of the data, and the resulting predictions of individual crew-member performance have been summarized by Anno et.al. (1984)

RESULTS

The data obtained from the questionnaire (Winne, et al., 1984) and the predicted effects on each crew position (Anno et al., 1984) indicate that the effects of radiation sickness are dependent on the physical demands of the task. This trend was tested across all observations using a 2-way ANOV to compare overall performance decrement for a crew positions with symptom severity. Newman-Keuls comparisons indicate that for all symptoms there is a significantly greater ($p<.01$) performance decrement for the physically demanding tasks (e.g., gun crew loader) than for the non-physically demanding tasks (e.g., fire direction center). Based on this finding, predicted performance decrement was divided into 30, 50, and 70 percent bands, with limits of each band determined by the decrement observed in the most and least affected task. This analysis, together with regions of no effect and incapacitation, are summarized in Figure 1.

Figure 1: Radiation-induced performance decrement as a function of time after exposure and radiation dose.



DISCUSSION

The results of this study clearly indicate that significant differences in the effect of radiation sickness on the performance of military tasks can be identified and quantified using a questionnaire. In order to determine whether these differences are representative of performance decrements which are observed in the actual performance of military personnel during illness, these results were compared with the performance measures taken on Naval and Coast Guard Personnel during seasickness (Wiker, et al., 1980). Matched for symptom complex and task, the results of the two studies agree. Likewise, the finding of this study agree with the findings in animal research that postirradiation performance decrement increases with task demand.

Two aspects of the effects of radiations on the battlefield were not directly addressed by the questionnaire study. These were the psychological aspects of nuclear combat and the transient incapacition which can be produced by radiation exposure. The psychological impact of nuclear combat on military effective is a very difficult problem unto itself. This study did not attempt to quantify the performance decrement from the psychological trauma of nuclear battle. We did, however, conduct a comprehensive review of what is known in this area, what the problems are and what potential impact they might have. That review will be summarized in another paper in this symposium (Sessions, 1984). Transient behavioral incapacitation after irradiation was quantified from behavioral studies with monkeys. Too little data on incapacitation in man exists to permit a time and dose dependent description of its occurrance. Since this particular reaction is independent of the signs and symptoms of radiation sickness and has been carefully studied in monkeys, the primate data were used to quantify that aspect of the behavioral impact of prompt total-body irradiation. Taken together, these various sources of data present a comprehensive picture of the impact of radiation on the combat performance capability in Army crews. As such, this work suggests that the utilization of this type of questionnaire approach can provide a signiifcant tool for gathering data on performance in hazardous environments.

REFERENCES

(Due to space limitations, references have been omitted from this edition of the paper. References are available upon request from Robert W. Young, Armed Forces Radiobiology Research Institute, Bethesda MD 20814.)

AD-P003 254

A Questionnaire Assessment of Estimated Radiation Effects Upon Military Task Performance

Albert S. Glickman, Peter S. Winne,
Ben B. Morgan, Jr., and Richard B. Moe
Organization Research Group of Tidewater, Inc.
Norfolk, Virginia

Abstract

One hundred twenty-five supervisors in four types of U.S. Army combat systems estimated the degree of degradation of military tasks for 30 descriptive symptom complexes associated with various radiation exposures. Results indicated that (a) the relative order of symptom effects were highly consistent across positions and the types of systems, (b) performances were expected to be deleteriously affected under most illness conditions, even mild ones, but incapacitation was not anticipated until illness conditions became quite severe, and (c) the most important factors in estimating performances were Fluid Loss and Fatigability/Weakness.

Introduction

The relationship between radiation sickness and human performance is largely speculative. A moderate amount of field data exists. But, in a recent review, Morgan and Cruser (1982) found only three studies dealing directly with the performance effects of irradiation. Anno, Brode and Washton-Brown (1982), concluded that there is not yet enough data available upon which to base operational estimates of the potential impairment of combat units by radiation sickness. The present study reports a method for obtaining estimates of radiation effects on combat performance from military personnel.

The vehicle for obtaining estimates was a questionnaire in which supervisors from four types of military crews made judgments of the effects of illness symptoms on task performance. These were the Artillery Gun, Fire Direction Center (FDC), Armor (Tank), and TOW-ITV System crews. The expert judges were given brief descriptions of sets of radiation sickness symptoms and asked to make estimates of how tasks critical to accomplishment of a team's mission would be affected. Ratings for specific tasks were combined into estimates of expected degradation for each position in each type of crew. A total of 15 different crew positions were investigated.

Method

U. S. Army combat crews were chosen because they are small, interdependent, cohesive groups whose effectiveness requires time-critical task performance that can be observed.

Analysis of data from 125 supervisors are reported here. A total of 21, 23, 40, and 41 personnel represented Artillery (105mm), FDC, Armor, and Mounted TOW, respectively. Across the four crews the mean time in service and specialty were 8.5 and 6.6 years respectively.

Survey Instruments

Four questionnaires, each containing the critical tasks for one crew, were prepared, consisting of: (a) a demographic information section; (b) introductory pages containing instructions, examples, and a nuclear combat scenario; (c) a sample response sheet; and (d) task performance estimates for 30 symptom sets. Each symptom page contained a description of a set of radiation sickness symptoms followed by a list of the selected military tasks performed by the given crew members.

These 30 symptom sets were drawn from six physiological components of illness: (a) upper gastro-intestinal (UG), (b) lower gastro-intestinal (LG), (c) fatigability/weakness (FW), (d) cardiovascular (CV), (e) bleeding and infection (BI), and (f) fluid loss and electrolyte imbalance (FL). Each of these contained five behavioral descriptions representing five points on a scale of severity. Written descriptions of sickness symptoms were formed by selecting one component symptom from each of the six syndromes. For example, a symptom coded "334231" combined severity level "3" of UG, "3" of LG, "4" of FW level, "2" CV, level "3" BI and level "1" of FL. Severities were associated with a wide range of radiation intensities and times after exposure.

For each task a referent time, derived from observation, technical manuals, and expert judgment, was listed, which represented usual completion time under normal (healthy) conditions. The respondents were instructed to select one of three alternatives describing the effect of the symptoms shown: (1) No effect, (2) Task would take longer to accomplish, or (3) Could not do it at all. For the second alternative, respondents were asked to estimate how long it would take for a crew member to complete the task.

Procedure

Four Army training sites were visited during November 1982 through February 1983. The questionnaire was administered to groups of 4 to 44. The sessions, approximately two and one-half hours in duration, started with a briefing on the nature, purpose and use of the data. Following this, instructions were given and a sample response page was "walked through" on a task-by-task basis. Respondents then completed the task lists for each of the 30 illness conditions. Ten sets of three complexes were formed; each set contained a relatively mild, moderate and severe complex. The order of presentation of these 10 sets was randomized in the questionnaire, and used for all subjects, except for a second sample from the FDC group, in which the order was reversed.

Results

The analyses of the questionnaires focused on three issues: (a) what was the degree of consistency in the estimates of the respondents; (b) what level of degradation was expected to result from the symptoms described; and (c) which of the six symptom components were most important in determining the expected decrements.

Measures

To estimate the effects of radiation, the data were treated in two ways. First, the percentages of respondents who checked "No Effect" and "Cannot Do" for each symptom set and task were computed. In several positions (e.g., Gunner, Assistant Gunner and Loader in the Artillery; HCO and Computer in the FDC; and the Gunner in the Tank and TOW), the inability to perform one task was grounds for classifying that crew member as incapacitated, because of the serial linkage of tasks.

Second, a percentage of baseline performance, based on the ratio of the reference time to the estimated time under the sick conditions, was calculated for each task. These percentages were further adjusted on a task-to-task basis by dividing each respondent's expected rating by his rating to symptom "112111" ("Somewhat tired with a mild weakness"). "No Effect" ratings were treated as 100% of baseline, and "Cannot Do" ratings were treated as 0%. Percentage of baseline measures for the tasks within each position were averaged. In turn, position means were averaged to derive crew-decrement estimates.

Rating Consistency

The first set of analyses was conducted to test the extent to which the expected severity of the symptom sets was consistent across all positions; i.e., did the judges agree on the relative magnitude of effect of the 30 symptoms on task performance? The 30 symptoms were rank-ordered on mean percentage of baseline performance for each position within each crew. The test of agreement among the rank orders, using the coefficient of concordance, Kendall's $\underline{W}$, showed substantial consistency in the rank orders across crew positions within crews. The average value of $\underline{W}$ was .97, accounting for 94% of the variance of rank order. A similar test on the order of symptom effects across crews also found high consistency, $\underline{W}=.94$, accounting for 88% of the variance. The implication is straightforward--that the relative severity of degradation in response to illness is expected to be similar across positions and crews, even though the particular tasks performed by crew members are different.

Expected Degradation

Evidence of rank order consistency, however, does not provide a test for the absolute degree of expected effect. Some positions might be more detrimentally affected by radiation sickness than other positions. Accordingly, the degradations expected for crews could differ. To assess the extent to which positions were expected to degrade as a function of illness conditions, the two measures of performance (e.g., percentage of respondees who reported the position would be incapacitated, and the mean percentage of baseline performance) were assessed within and across the symptom set. Comparisons were made both among the positions in each crew and among crews.

It was found that performance would be affected somewhat under most illness conditions, even relatively mild ones, but that incapacitation was seldom expected until relatively severe conditions prevailed. Across positions and symptoms, 27% of the respondents gave "No Effect" responses and 13% gave "Cannot Do" responses.

At milder levels of illness, only a small percentage of respondents felt that performance would be incapacitated, regardless of position. At moderate and severe levels, 20% to 60% of the respondents expected that the requirements

of positions could not be performed. Under the most severe illness condition, more than half (55%) of the respondents expected incapacitation to occur.

The analysis of percentage of baseline data provides additional information regarding expected decrement level. Figure 1 shows the mean percentage of baseline performance for each of the crews as a function of symptom severity. Across symptom sets, the mean expected degradation was 64%. With the mildest symptom ("113111"), mean expected performance across crews was 93% of baseline. For the most severe symptom ("415314"), the average rating was 24%. The trend of degradation across symptoms was generally linear. The FDC crew was judged to be somewhat less degraded than the other three crews, probably because of the less physically demanding nature of their jobs.

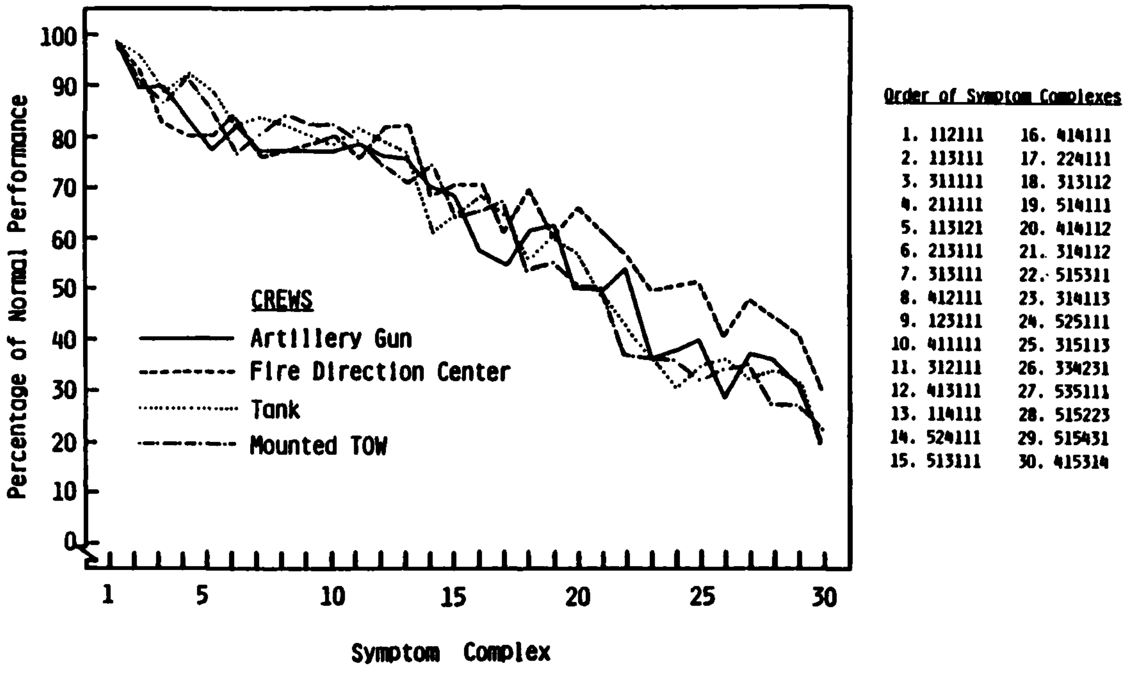


Figure 1. Mean percentage of normal performance for crews as a function of perceived symptom complex severity across the four crews.

Within position, the standard deviations suggest that there were large individual differences in estimates. For this reason, the absolute values are judged to be less important then relative ranking of the illness conditions. Nevertheless, the respondents in all groups expected a fairly steep decline in performance as a function of illness. Except for the mildest symptoms, performances at all positions were expected to be 80% of normal or worse. In fact, mean expected performance for nearly one-half of the symptoms used was less than 60% of normal. Physically demanding positions were judged to be most vulnerable to radiation effects.

Symptom Components

Finally, stepwise multiple linear regression was used to determine the amount of variance in the degradation estimates accounted for by the six

physiological components, and to identify the components with the highest beta weights, within and across all positions.

Across positions, the most important symptom components were Fluid Loss, (characterized by dry mouth and headaches, and at high scale levels, burning sensations and difficulty moving), and Fatigability/Weakness. The positions requiring physical strength were somewhat more dependent on the degree of FW associated with the illnesses than other positions. The final regression equation in the combined analysis accounted for approximately 41% of the variability in the decrements expected across the symptom sets, suggesting that it is feasible to predict to other illness outcomes from knowledge of the symptom components.

Discussion

This study offers only a beginning in our understanding of the expected effects of radiation illness, but several of the findings are rather robust. Strong consistency across both positions and crews in the order of severity of the symptoms is shown. The range of tasks and skills tapped suggests that declines in performances will be general and not limited to a few jobs.

The findings from the regression analysis would justify further investigation. The results of this preliminary study were encouraging with respect to prediction of expected performances under other sickness conditions. The weights derived for the symptom components provide one means for predicting dose/time curves of military performances. At this point it appears that Fluid Loss and Fatigability/Weakness will consistently result in the most degradation.

Finally, the level of decrement data, in terms of both the percentage of "No effect" responses, and the percentage of baseline performance, reflects a general expectation of systematic degradation for all of the positions sampled, even at mild symptom levels.

We close with these caveats. This is a study of judgments, not of actual experience, and only physiological factors have been dealt with here. Because other factors are present in a nuclear engagement, such as blast damage, equipment failure, and psychological reactions, our estimates of deleterious effects of exposure to radiation are very conservative.

References

Anno, G. H., Brode, H. L., and Washton-Brown, R. (1982). Initial human response to nuclear radiation (PSR Note 477). Los Angeles, CA: Pacific-Sierra Research Corporation.

Glickman, A. S., Winne, P. S., Morgan, B. B., Jr., Moe, R. B. (1983). Estimated effects of intermediate levels of nuclear radiation upon the performance of military tasks: A questionnaire assessment (Report No. 82-2/3-1). Norfolk, VA: Organization Research Group of Tidewater, Inc.

Morgan, B. B., Jr., and Cruser, L. G. (1982). Performance effects of sickness and symptomatology: A review of the literature (Report No. 82-1-2). Norfolk, VA: Organization Research Group of Tidewater, Inc.

Estimated Effects of Ionizing Radiation Upon Military Task Performance: Individual Combat Crewmember Assessments

George H. Anno and Don B. Wilson
Pacific-Sierra Research Corporation
Los Angeles, California

Abstract

Quantitative estimates are developed of the performance levels for selected individual Army combat crewmembers exposed to prompt ionizing radiation from nuclear weapons. The performance levels, expressed in percent of normal (baseline) task performance, provide information for military operations planning, combat training, and computer simulation modeling of combat crew and unit effectiveness. The methodology is described where data from two separate bodies of information:- 1) acute radiation sickness symptomatology, and 2) judgment of task performance time from Army combat crew questionnaires - are integrated to compute performance levels as a function of dose (free-in-air) and post-exposure time.

Introduction

The procedural methodology for developing the dose/time individual performance relationships described utilizes information resulting from preceding work performed for the DNA Intermediate Dose Program effort including a comprehensive review and assessment of acute radiation effects/symptomatology in humans [Baum, et al., 1983]; description and quantification of radiation sickness in terms of five levels of severity, and aggregation of symptom groups into symptom complexes [Anno, et al., 1983]; and gathering judgmental data on military task performance from questionnaires administered to Army combat crew personnel [Glickman, et al., 1983].

The acute radiation effects and symptomatology descriptions by Baum, et al. for each of eight (8) separate dose ranges covering exposures from 75 to 4500 rads (cGy), provides the medical/radiobiological basis for establishing the five symptom severity levels and forming the symptom complexes. Symptom complex verbal descriptions used in the Army combat crew questionnaires form the link between dose/time and the task performance estimates provided by the combat crewmembers.

We focus on crewmembers representing four different crews specifically selected for study because their operations include the prospect of decisive encounters with the enemy. The term "decisive" means that one of the combatant crews will probably be put out of action as a consequence of the encounter. The crews are small, with three or four members, and the encounters are usually brief, typically lasting between 30 and 90 sec. Accordingly, the crew missions, operational scenarios, and crewmembers' tasks can be readily modeled.

Analysis

Performance data, derived from questionnaires administered to a total of 161 senior noncommissioned officers from the four combat crews were evaluated to identify those applicable for further analysis based on criteria imposed. Twenty four (24) respondents were excluded based on either lack of experience, background, self-disqualification or missing questionnaires. Another 21 were rejected because of noncompliance with instructions or unreasonable response patterns, leaving a total of 116 or 73 percent. The table below summarizes combat crew information used for the regression analysis.

Combat Crew Summary for Regression Analysis

Combat Crew	Number of Respondents (N)	Number of Symptom Complexes (K)	Crewmember Position	Number of Tasks (J)
Artillery, Gun (155 mm Self-Propelled Howitzer)	25	30	Chief-of-Section	1
			Gunner	2
			Assistant Gunner	2
			Loader	6
Artillery, Fire Direction Center	21	30	Fire Direction Officer	3
			Horizontal Control Operator	2
			Computer	2
Armor, Tank (M60A3)	34	30	Tank Commander	4
			Gunner	3
			Loader	2
			Driver	1
Anti-Armor, Mounted TOW (TOW/ITV)	36	40	Squad Leader	2
			Gunner	2
			Driver	1
			Loader	2

A linear regression analysis was performed of respondent-averaged data to obtain the fit parameters for predicting the performance for each of the 15 crewmember positions. The logit form was used to model the response data since it guarantees that all predicted performance values including the upper and lower confidence limits lie within the interval (0,1) which is consistent with the dependent variable (performance) input data. That is,

$$\ell n \; \frac{P_k}{1 - P_k} = \underset{\sim}{x}_k' \, \underset{\sim}{\beta} + \varepsilon_k \quad , \quad 0 < P_k < 1$$

where,

$$P_k = \frac{\sum_{j=1}^{J} t_{oj}}{\sum_{j=1}^{J} \left(\frac{t_{oj}}{\bar{P}_{jk}} \right)} \; , \quad \bar{P}_{jk} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{t_{oj}}{t_{ijk}} \right) \; ,$$

and where P is performance, t is estimated time to complete a task, t_o is reference or baseline time to complete a task, and i,j, and k are, respectively, respondent, task, and symptom complex indices. $\underset{\sim}{x}'_k = (1, x_{k1}, x_{k2}, \ldots, x_{k6})$ is a vector of integers $(1 \leq x_{k\ell} \leq 5)$ denoting symptom severity level for the k^{th} symptom complex and ℓ^{th} symptom category; the six symptom categories whose severity levels were considered in this analysis are 1) upper gastrointestinal distress, 2) lower gastrointestinal distress, 3) fatigability/weakness, 4) hypotension, 5) infection/bleeding, and 6) fluid loss/electrolyte imbalance. The β-parameters are given by the vector $\underset{\sim}{\beta}' = (\beta_0, \beta_1, \ldots, \beta_6)$ which are estimated. The relationship for estimating performance of a crewmember for the k^{th} symptom complex is $\hat{P}_k = 1/[1 + \exp(-\underset{\sim}{x}'_k \hat{\underset{\sim}{\beta}})]$.

Figure 1 gives an example of predicted performance (straight diagonal line) and data from the questionnaire (asterisks). Also shown are 95 percent confidence bounds for the regression line (solid lines) and the data points (dashed lines). The regression analysis was performed using SAS [Brelsford, 1981]; STATLIB [Ray, 1982] was also used to make estimates of the correlation (ρ) of error terms utilizing maximum likelihood techniques; values of ρ = 0.3 to 0.4 were estimated for the data which suggests that the responses are correlated. Therefore, the confidence bounds may be from 12 to 30 percent larger.

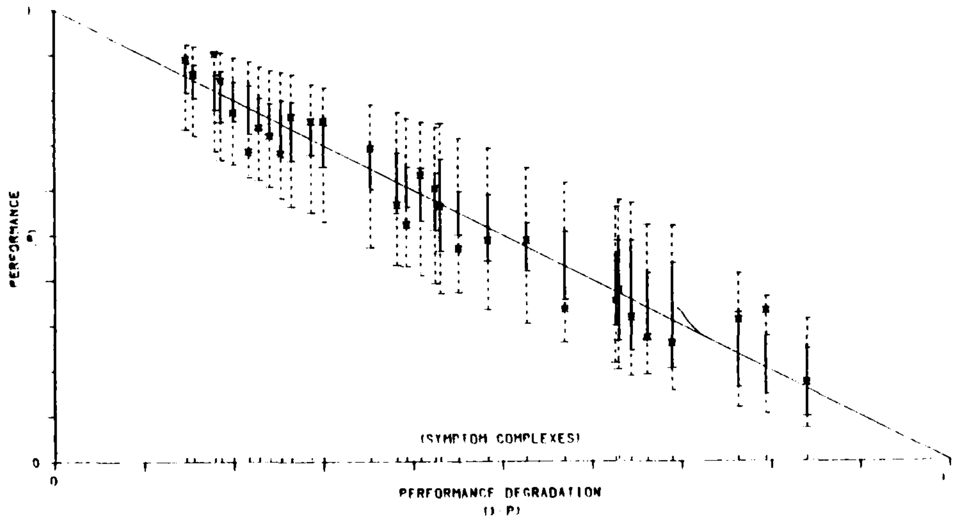


Figure 1. Tank crew individual performance: Tank Commander.

The performance predictions are mapped onto the dose/time plane applying numerical techniques to a framework of symptom complexes. The dose/time plane is divided into eight dose bands (75-150, 150-300, 300-530, 530-830, 830-1100, 1100-1500, 1500-2000, 2000-3000, and 3000-4500 rads (cGy)); each dose range in turn is divided into a series of adjacent finite increments along postexposure time to represent the dose/time location of acute radiation illness states defined by symptom complexes. The symptom complexes are each identified by a row of six (6) integers varying from 1 to 5 depending upon severity level; e.g., 213111 designates severity level 3 for upper gastrointestinal distress, severity level 1 for lower gastrointestinal distress, severity level 3 for fatigability/weakness, and severity level 1 for hypotension, infection/bleeding, and fluid loss/electrolyte imbalance.

Computational Procedure

In order to create a three-dimensional representation of performance as a function of dose and time, the symptom map described above plus imposed boundary conditions are utilized to build a smooth surface that both covers the regions of interest and passes through each performance input data point. To do this, a "french curve" algorithm is applied to translate the input data into an evenly spaced (logarithmic) grid of performance predictions for 18 values of dose and 39 values of time. The particular dose/time grid structure chosen allows full resolution of the underlying symptom map without adding an excessive amount of easily interpolated information.

Figure 2 is an example of the three-dimensional performance surfaces vs. dose and time shown for the tank crew commander. A similar numerical procedure was used to generate isoperformance contour plots which correspond to 10 percent incremental co-planar projections of the performance surface onto the dose/time plane; Fig. 3 shows the contours for the tank crew commander.

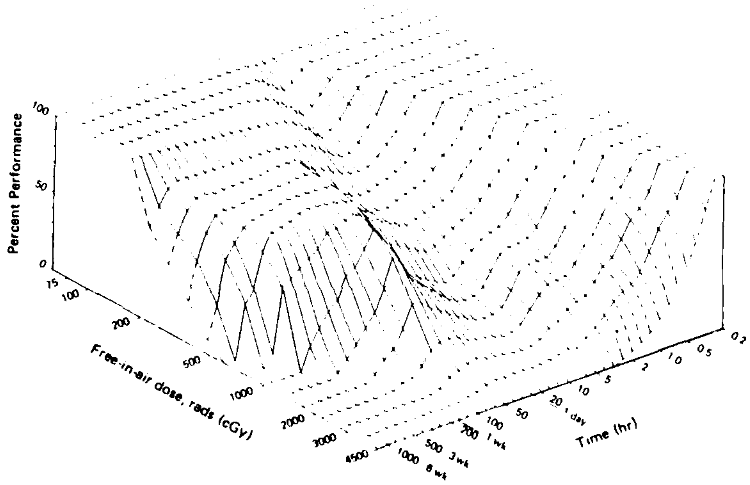


Figure 2. Performance surface: tank crew, Tank Commander.

Discussion

Performance vs. dose level and time after exposure such as illustrated in Figs. 2 and 3 provide the basic information for further developing and expressing the acute effects of ionizing radiation on combat troop effectiveness for specific applications. These include Army field manuals, instruction and training materials, and computer simulation of combat engagements. Fig. 2 provides an overall illustration of individual performance as related to dose (75-4500 rads (cGy)) and time (minutes to 6 weeks). Surface depressions show where performance is degraded. The initial overall depression is consistent with the prodromal symptomatology from medical observations which includes nausea, vomiting, fatigue, weakness, and initial fluid loss. For later times the broad rise in performance between 10 and 200 hrs in Fig. 2 is related to some remission of the prodromal symptoms prior to the onset of delayed symptoms such as infection, bleeding, fever, lower gastrointestinal distress, fluid loss and

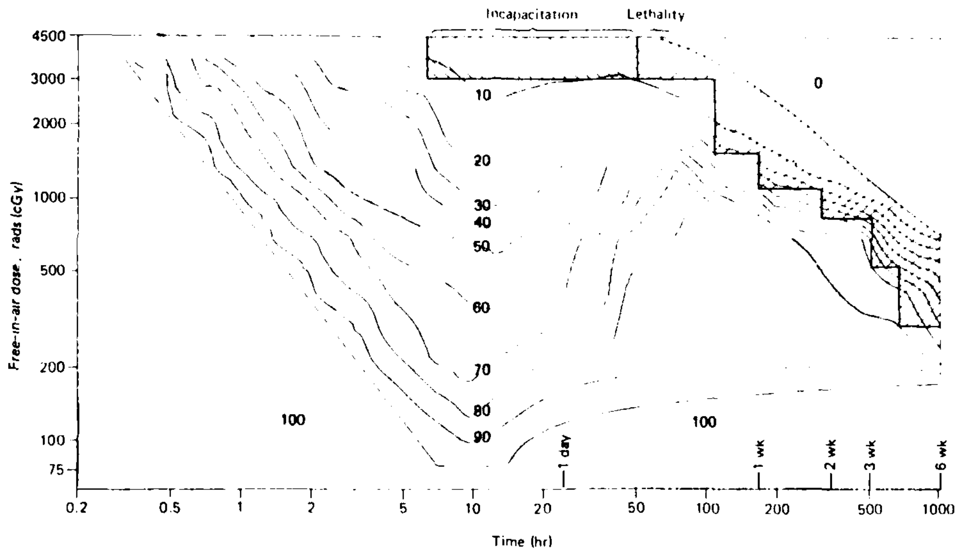


Figure 3. Performance contours, 10 percent performance intervals: tank crew, Tank Commander.

reoccurrences of nausea, vomiting, fatigability, and weakness, reflected by the downward slope of the surface. The broad rise decreases with increasing dose, which reflects the increasing absence of remission of radiation induced effects affecting performance. Medical experience suggests symptoms associated with acute radiation sickness affecting performance are present in those exposed as follows: 5-30 percent from 75 to 150 rads (cGy); 30-60 percent from 150-300 rads (cGy); 60-90 percent from 300-530 rads (cGy); 90-100 percent from 530-830 rads (cGy); and 100 percent from 830 rads (cGy) and greater.

References

Anno, G. H., D. B. Wilson, R. W. Young, B. B. Morgan, Jr., S. J. Baum and H. R. Withers, *Levels of Severity and Symptom Complexes for Acute Radiation Sickness: Description and Quantification,* (in progress).

Baum, S. J., G. H. Anno, R. W. Young and H. R. Withers, *Symptomatology of Acute Radiation Effects in Humans AFter Exposure to Doses of 75 to 4500 Rads (cGy) Free-in-Air,* Pacific-Sierra Research Corporation, Los Angeles, CA, Note 581, June 1983.

Brelsford, W. M. and D. A. Relles, *STATLIB--A Statistical Computing Library,* Prentice-Hall, Englewood Cliffs, NJ, 1981

Glickman, A. S., et al., *Estimated Effects of Intermediate Levels of Nuclear Radiation Upon the Performance of Military Tasks: A Questionnaire Assessment,* Organizational Research Group of Tidewater, Inc., Norfolk, VA, Technical Report 82-2/3-1, March 1983.

Ray, A. A., et al., (eds.), *SAS Users' Guide: Statistics,* SAS Institute, Cary, NC, 1982.

AD-P003 256

A Summary of the Psychological Effects of Tactical Nuclear Warfare

C. Rufus Sessions
U.S. Air Force Academy
Colorado Springs, CO

Abstract

The psychological component of the response of combat troops to tactical nuclear warfare is a troublesome variable which plagues military planners and commanders responsible for the preparation of the armed forces for the eventuality that nuclear weapons might one day be used in armed conflict. The devastating physical effects of nuclear weapons have been extensively documented and the biological response of animals and humans to radiation has also been well studied, but very little is known concerning the probable effects of experiencing a nuclear attack upon the emotional stability, morale, and motivation of soldiers to perform their assigned duties. The literature addressing this topic has been reviewed and evaluated as part of the Defense Nuclear Agency's Intermediate Dose Program. A summary of the report resulting from this effort is presented herein. The full literature review and accompanying bibliography are available from the author upon request.

Introduction

Any attempt at predicting the response of combat personnel to the threats associated with tactical nuclear warfare must deal with two separate but interrelated problems. One concerns the assessment of the physical and mental performance capabilities of individuals who have experienced a nuclear weapons attack; that is, the capacity of individuals to function and carry out their normal duties following exposure to the range of doses of ionizing radiation expected in the tactical nuclear environment. The second problem concerns the assessment of the emotional stability and willingness of individuals to perform their duties following an attack given that they possess the capability. Even under the best of circumstances, given good leadership and training, the combat effectiveness of soldiers is highly correlated with their emotional preparedness for battle, individual and unit moral, and unit cohesiveness. Given that exposure to the effects of tactical nuclear weapons, or even the threat of exposure, can be expected to produce levels of psychological stress which could result in degradation of these factors, the psychological effects of tactical nuclear warfare must be considered an important variable

affecting the response of soldiers to intermediate doses of ionizing radiation.

The primary efforts associated with the intermediate dose program have focused on the capacity of individuals and crews to function following exposure to radiation, primarily because the nature of this problem is more clearly understood and more directly addressable through empirical investigation. Although the psychological factors will almost certainly affect the performance of soldiers on the nuclear battlefield, the extent of this effect cannot be accurately estimated on the basis of current knowledge. In fact, this problem cannot be addressed quantitatively because there are no hard data which would permit an objective estimate of the manner in which performance would be affected. Despite this difficulty, discussion of the psychological issues is important in order to (a) place the estimates of performance capabilities in the context of the broader perspectives of the realities of combat, and (b) provide a framework within which future research efforts in this area may be developed.

The available information directly relevant to this problem is very limited. The only two instances of the use of nuclear weapons in warfare at Hiroshima and Nagasaki involved primarily civilian populations and provide little information directly relevant to the response of soldiers and their residual combat capability. Therefore, indirect sources of information were considered in attempting to estimate the psychological consequences of tactical nuclear combat and their effects on combat performance. Some of the sources providing information pertinent to this problem included those that address conventional combat psychiatry, World War I gas warfare experiences, human stress, medical psychology, civilian disasters, the Three Mile Island experience, and human shock or grief responses.

Summary

The psychological component of the stress impinging upon the soldier facing the threat of the modern integrated battlefield has often been addressed with the most attention being given to the general deleterious effects of combat stress, fear, fatigue, and most recently, exposure to chemical weapons. The integrated battlefield concept emphasizes the combination threat of weapon systems available in the modern military arsenal, including conventional small arms, high explosive artillery, tactical air support ordinance, as well as chemical, biological, and nuclear weaponry. The tactical nuclear combat environment is a subset of the total combat environment of the modern integrated battlefield. The psychological effects of the consequences of tactical nuclear warfare has received less attention than conventional or chemical warfare, perhaps because of a) the lack of historical experience with nuclear weaponry, or b) the inability to realistically conceive the threat, or, the human tendency to deny or repress

concepts so threatening and overpowering, or c) perhaps due to the general lack of attention to nuclear weaponry during the recent political climate of detante and antinuclear sentiment.

Previous efforts to address this problem have focused upon estimating the occurrence of extreme emotional reactions leading to panic behavior or neuropsychiatric casualties resulting from total emotional breakdown. The former response has been generally ruled out as an expected occurrence (Glass, 1956). The latter response has been examined thoroughly based upon the voluminous data bases from the previous combat experiences in World War II and Korea (Vineberg, 1965).

A straight forward recognition of the fact that intensity of combat directly correlates with NP casualty rates leads to the conclusion that casualties due to emotional breakdown will be higher in combat involving nuclear weapons than any previous conflicts involving conventional weapons alone. Estimates of neuropsychiatric casualty rates based on ratios referenced to killed or wounded in action are likely to be misleading, due to the fact that much higher physical casualty rates will be expected from the employment of tactical nuclear weapons.

The greatest impact on residual combat capability in tactical nuclear combat may be expected to derive not from neuropsychiatric casualty rates, but from emotional disruptions which are debilitating from a performance point of view but not severe enough to produce the classical picture of emotional breakdown represented in the neuropsychiatric casualty.

The following description of the emotional disruptions expected from troops exposed to nuclear attack are based on materials describing or analyzing the behavior of survivors of the nuclear attacks on Hiroshima and Nagasaki in 1945. These materials include eyewitness accounts Siemes, 1946; Hachiya, 1955), edited reports of eyewitness accounts (Hersey, 1946; Nagai, 1951), and studies involving interviews with survivors of the bombings (Janis, 1951; Kubo, 1952; Lifton, 1967). While the relevance of the behaviors of the civilian populations of these two Japanese cities in 1945 to the expected behavior of modern combat troops must be continually questioned, this body of information represents the only empirical evidence regarding the psychological response of human beings to nuclear attack and therefore should be carefully considered.

The experience of being exposed to nuclear attack can be expected, in the short run, to be an intensely traumatic emotional stimulus, capable of rendering individuals temporarily ineffective in performing routine military duties. While panic behaviors and long term emotional disturbances may not be expected, a state of temporary emotional disturbance is to be expected in a sizeable but unknown proportion of individuals surviving nuclear attack, characterized by a shocked, dulled affect, with little capability

for behavior beyond that required for immediate survival and escape from the perceived source of danger. Attempts at other behaviors may be ineffectual or inappropriate, with the exception of a small percentage of cases in which heroic, goal-oriented, selfless behavior may be expected. Depending upon the combination of many variables affecting the severity of such emotional disturbances, this situation may be expected to persist for a period of hours to days in the majority of cases.

The variables affecting the susceptibility to and severity of emotional disruptions include training and preparedness, combat experience, leadership, unit cohesion, perceived degree of danger, degree of chaos and destruction in the immediate environment, and degree of isolation. Variables affecting the rapidity of recovery include safety and rest, training and preparedness, leadership in the immediate post-trauma period, cohesiveness and group support, and immediacy, expectancy, and proximity of psychiatric first aid treatment.

The behavior of individuals following nuclear attack can best be analyzed within a model that takes into account the dimensions of time and space (Logan and Killian, 1953). Since the emotional responses to disastrous events are characterized as fluid, changeable, and transient (Glass, 1956), the expected behavior of affected individuals must be referenced to a time sequence following the traumatic event. Likewise, because the physical destruction and damage resulting from a nuclear explosion extends in a diminishing fashion over a wide area, the behavior of individuals reacting to the destruction will be impacted inversely proportionately to their distance from the epicenter.

Recommendations

The most important factor affecting the psychological effects of nuclear warfare that is under the control of military command is the training and preparedness of individuals and leaders to cope with the trauma of nuclear combat. Emphasis in training should be on a) realism, in order to reduce the shock impact of the aftermath of nuclear attack, b) accurate information concerning the threat that is faced, and c) information limited to that which is readily comprehended and assimilated by the average individual and which is directed towards the promotion of his or her personal welfare. New research is recommended to determine the type and amount of information best suited for the training of today's combat troops.

References

Glass, A. J. (1956). Psychological considerations in atomic warfare. *U.S. Armed Forces Medical Journal*, *7*, 625-639.

Hachiya, M. (1955). *Hiroshima diary* (W. Wells, Trans.). Chapel Hill: University of North Carolina Press.

Hersey, J. R. (1946). *Hiroshima*. New York: Alfred A. Knopf.

Janis, I. L. (1951). *Air war and emotional stress* (1st ed.), (Psychological studies of bombing and civilian defense). The RAND Corporation. New York: McGraw-Hill.

Kubo Yoshitoshi. (1952). Study of human behavior immediately after the atomic bombing of Hiroshima: A socio-psychological study of the atomic bomb and atomic energy. I. [Eng. abstract]. *Japanese Journal of Psychology*, *22*, 103-110.

Lifton, R. J. (1967). *Death in life: Survivors of Hiroshima*. New York: Random House.

Logan, L., & Killian, L. M. (1953). *Troop reaction to atomic attack-A preview* (ORO-T-205). Chevy Chase, MD: Johns Hopkins University, Operations Research Office. DTIC

Nagai, T. (1969). *We of Nagasaki* (I. Shirato & H. Silverman, Trans.). New York: Meredith Press. (Original work published 1951 by Duell, Sloan & Pearce).

Sessions, G. R. DNA Intermediate Dose Program Report: The Psychological Effects of Tactical Nuclear Warfare, (In Preparation).

Sessions, G. R. Psychological Effects of Tactical Nuclear Warfare: A Bibliography, (In Preparation).

Siemes, J. A. (1946). Hiroshima-August 6, 1945. *Bulletin of the Atomic Scientists*, *1*, 2-6.

Vineberg, R. (1965). *Human factors in tactical nuclear combat* (technical report 65-2, contract DA 44-188-ARO-2). Alexandria, VA: George Washington University, Human Resources Research Office.

AD-P003 257

PHYSICAL PERFORMANCE TESTS AS PREDICTORS OF TASK PERFORMARCE

T.L. Doolittle Ph.D.
University of Washington

Oscar L. Spurlin Ph.D., and M. Peter Scontrino Ph.D.
Ergometrics and Applied Personnel Research Inc.

Abstract of Panel Presentation

Test Development:

The more arduous the task, the greater the intensity of force which must be applied per unit of time to overcome resistance or achieve rate. Intensity is commonly called "workload" with magnitude expressed in appropriate units of power. Two complex factors determine the limits for which an individual can produce energy and generate the requisite power:

1. capacity to utilize oxygen, and
2. ability to generate muscular tension.

The former is called "aerobic power" and the latter "strength". For repetitive tasks, over an 8 hour workday (with normal breaks), individuals normally will not function at an intensity greater than 40% of their maximum aerobic power and/or 15% of maximum strength. The limit for occasional lifting, while influenced by posture and body mechanics, normally should not exceed 70% and never exceed 90% of the individual's maximum strength.

Aerobic power requirements are estimated from heart rate response to known and unknown workloads. From rest to approximately 70-80 percent of maximum aerobic power, increases in heart rate correlate linearly with increases in energy expenditure and oxygen consumption (Figure 1).

Further, since the heart can beat only so fast and still function as an effective pump, an individual's maximum functional capacity can be projected from the heart rate response at a given sub-maximum workload. Inter-individual comparisons also are possible. In Figure 1, a given workload (dashed line) is relatively more demanding for person "A" than for person "B", and since both will "max"at about the same heart rate, "B" has more reserve and should be capable of a higher maximum workload. Once the baseline relationship between heart rate and workload has been established for each individual, heart rate on job tasks (dotted lines in Figure 1) enables one to accurately estimate the metabolic requirements. This protocol has been utilized and is widely reported for a variety of occupational endeavors in the literature (Astrand, 1977).

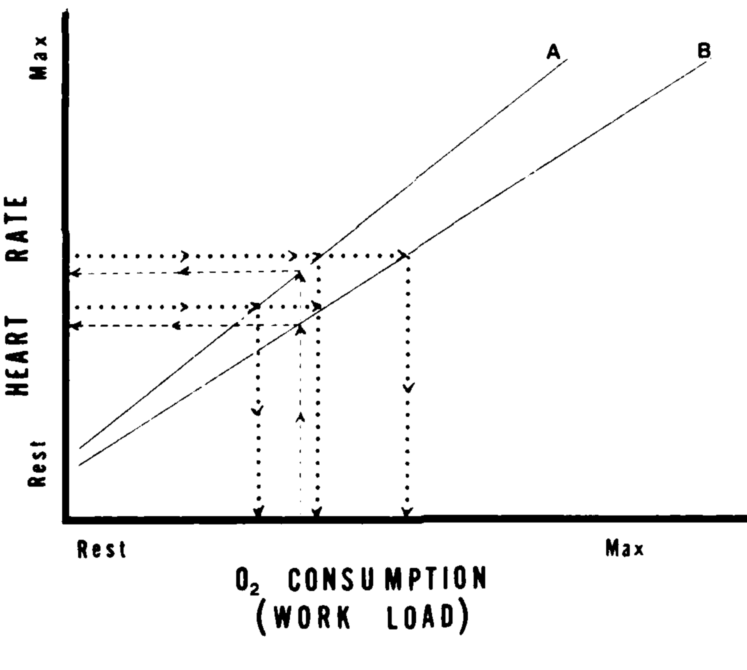
Strength requirements are established by determining the forces required to lift, push, pull, shovel, slide, etc. as appropriate. Multiple samples are necessary to account for variability. Heights and distances the objects must be moved, and total amount of material per unit of time must be obtained. From these data strength requirements are determined and expressed in terms of pounds per lift (for occasional endeavors) or foot pounds per minute for repetitive tasks. While the aerobic power requirements can be generalized

for measurement on a treadmill, cycle ergometer, or bench step, strength testing should closely replicate the task to insure that the requisite muscle groups are being tested.

For occasional lifting, requiring the individual to lift an object of similar weight five (5) times will insure that it does not exceed approximately 85% of capacity. For repetitive strength tasks 67% of capacity can be performed at 20 contractions per minute for short durations. Thus, for a workload requirement of 750 ft-lb/min, there are two defensible alternatives:

- a single maximum test requiring 6250 ft-lb/min, or
- a three to five minute effort at 4200 ft-lb/min, accomplished at a rate of 20 repetitions per minute.

FIGURE-1 Heart Rate & Work Load



<u>Validity Evidence and Impact on Protected Groups.</u>

From the foregoing discussion it can be seen that it is possible to replicate the significant components of physically demanding occupations. If a test can be demonstrated to represent important job components it is valid to use the test in applications such as pre-employment screening. Nevertheless, because of the legal guidelines and changing professional standards surrounding test validation, there are some important issues to consider in order to firmly establish the defensibility of a physical performance test.

Content vs. criterion-related validity

The primary issue here is whether job analysis alone can provide a sufficient basis for validation. Our experience is that it is, in fact, a superior method. Job analysis procedures as described above, provide a detailed description of actual task requirements. This description is independent of the characteristics of current incumbents, it is not dependent on large numbers of employees performing a particular job, and it is not dependent upon supervisory ratings or other highly subjective methods.

The difficulty with criterion-related validation studies is the criterion. Physical performance tests are often validated through use of work samples as performance criteria. Examples can be given to show that the subjectivity and potential for measurement error is greater in this process, than if the entire research effort was concentrated on quantifying actual work tasks and replicating the critical components of these activities.

An issue for both these approaches is which tasks are truly critical to job performance. Is a task that is performed twice a year a sufficient justification for a screening standard? Emergency service personnel are often screened on the basis of worst-case scenarios. This type of rationale requires substantially more justification. Another frequently confronted case is the 'labor pool' where only a certain proportion of individuals are ever assigned the arduous tasks.

Simulation

Perhaps one the more difficult questions concerns the degree of 'true-to-life' simulation necessary to be assured of a job-related test. That is, does a test need to involve exactly the same muscle groups used in exactly the same range of motion to provide an accurate prediction of job success.

In cases where the objective is to select among a relatively large number of applicants for a given opening, a high degree of precision is usually not necessary. This finding results from the moderate to high correlation between different types of physical performance tests. Leg strength is different from trunk strength, and stamina is different from both; yet tests of these dimensions might correlate on the order of r=.40. Thus, a typical finding would be that a single test can often provide roughly the same rank ordering of candidates as an entire battery of tests.

In situations that involve the use of highly developed physical abilities, then simulations are more appropriate in order to assess specific job-related capabilities that it would be impractical to learn or develop on the job. Rehabilitation therapy is another instance where more detailed knowledge of specific physical abilities may be needed. We have seen a real need for return-to-work standards that are based on objective job related measures of capability.

Impact on protected groups.

Physical performance tests will have an impact on women and to a lesser extent other minorities of smaller average stature. There are, of course

real population differences in strength and endurance, and tests will reject a higher proportion of females than males. This does not mean that physical performance tests are automatically subject to successful legal challenges.

It is reasonable and legal to require job applicants to possess the qualifications necessary to perform the job effectively and safely. Physical performance tests will be well accepted by all concerned, provided:

- Job-relatedness is shown through research that conforms to legal and professional standards

- There are no alternative tests or procedures which could identify qualified candidates without the same impact.

- The job itself cannot be easily redesigned to eliminate the physically limiting tasks.

- The tests are used to screen 'in' physically qualified females and minorities.

In practical terms, a hiring agency simply needs to recruit and screen more female applicants to insure the placement of enough qualified and potentially successful individuals.

Utility: The Bottom Line

Over thirty years ago Brogden (1949) used linear regression to demonstrate the relationships of cost of selection, validity, and the selection ratio to the utility of a test. In this context utility refers to the dollar savings to an organization resulting from the higher performance (improved productivity) of those employees selected using a validated test. The higher the average test score of those selected, the greater the utility of the test.

Recently Schmidt et. al. (1979, 1982, 1983) and Cascio (1982) have revived interest in utility theory by providing inexpensive and straightforward procedures for estimating the utility of a selection procedure. Cascio (1982) has extended this process to assessing the utility of a wide range of human resource programs in organizations.

The basic formula for estimating the utility (the total improvement in performance over random selection) of a test as follows:

$$\Delta\mu = N_s r_{xy} SD_y \lambda / \phi - N_s C_y / \phi \qquad (1)$$

where:

$\Delta\mu$ = total gain over random selection
N_s = number of applicants selected
r_{xy} = validity of the predictor, when evaluated against the dollar-valued job performance criterion
SD_y = standard deviation of the dollar-valued job performance criterion
λ = the ordinate (height) of the normal curve corresponding to the predictor cutoff (the selection ratio)
ϕ = the selection ratio
C_y = cost of putting one person through the selection process

If one wishes to compute the gain in utility from substituting one selection procedure for another, the formula is:

$$\Delta \mu = N_s (r_1 - r_2) SD_y \lambda / \Phi - N_s (c_1 - c_2) / \Phi \qquad (2)$$

Where all terms are as defined above, except that r_1= validity of the new procedure, C_1= cost of the new procedure, C_2= cost of the old procedure.

In these formulas the one piece of information that is not readily available is the SD_y. Both Schmidt and Cascio present simple methods to estimate this dollar valued variance in performance.

Of course, physical performance testing yields an equal or larger benefit in terms of cost savings resulting from decreased lost time injuries. Using the above procedure and including the safety related benefits has led us to conclude that physical performance testing can lead to an increase in value of 15-30% for each individual hired. This translates into $2000 - $3000 per employee annually in a typical application.

REFERENCES

Astrand, P.O., & Rodahl, K. Textbook of Work Physiology (2nd Ed.). McGraw-Hill, New York, 1977.

Brogden, H.E. When testing pays off. Personnel Psychology , 1949, 2, 171-183.

Cascio, W.F. Applied Psychology in Personnel Management . Reston, VA: Reston Publishing Company, 1982.

Schmidt, F.L., Hunter, J.E., McKenzie, R.C., & Muldrow, T.W. Impact of valid selection procedures on work-force productivity. Journal of Applied Psychology , 1979, 64, 609-626.

Schmidt, F.L., Hunter, J.E., & Pearlman, K. Assessing the economic impact of personnel programs on workforce productivity. Personnel Psychology , 1982, 35, 333-347.

Schmidt, F.L., Hunter, J.E. Quantifying the economic impact of psychological programs on workforce productivity. Amer. Psychologist ,1983,38,473-478

PANEL SESSION

SIMULATOR-INDUCED SICKNESS: REACTION TO A TRANSFORMED PERCEPTUAL WORLD

Lawrence H. Frank and Gilbert L. Ricard
Naval Training Equipment Center

Robert T. Hennessy
National Research Council

Robert S. Kennedy and Michael E. McCauley
Essex Corporation

Herschel W. Leibowitz
Center For the Advanced Study in the Behavioral Sciences

Background

There have been numerous recent documented and anecdotal reports of aircrews experiencing psychophysiological disturbances, visual illusions and sickness following the use of flight simulators. Symptoms of simulator sickness occur not only during simulator flight, but, in some individuals, have lasted up to several hours post exposure. Furthermore, simulator aftereffects may be delayed; some aircrews report symptom onset as late as eight to ten hours post utilization. Incidents of simulator sickness have been documented in fighter, attack, patrol and helicopter simulators. These occurrences have been reported in both motion-base and fixed-base simulators, to pilots and other aircrewmen, as well as instructors. Preliminary data suggest that more experienced aircrewmen are at greater risk and that such factors as wide field of view and visual/inhertial lag contribute to the problem. Simulator sickness represents a major obstacle to obtaining the full training potential from the vast inventory of flight simulators currently in use and under development. Obviously the learning capability of an individual who is suffering discomfort generated by a simulator is greatly compromised. Moreover, there is the possibility that the visual and proprioceptive cues responsible for simulator sickness may contribute to negative transfer of training in actual aircraft flight.

This evidence suggests that simulator-induced sickness symptomatology resembles motion sickness and other forms of distress which occur after exposure to altered and rearranged sensory information.

Purpose

The purpose of this proposed panel discussion is to disseminate to a broad-spectrum audience what is known about simulator-induced sickness: its causes, its implications, countermeasures and recommended research issues.

Recently the import of simulator-induced sickness was highlighted by a National Academy of Sciences, National Research Council, Committee on Human Factors sponsored workshop on this topic. It was evident from the workshop that simulator sickness is of interest to many diciplines (e.g., sensory psychology, physiology, human factors, perceptual psychology, etc.). This proposed panel discussion program will include five of the invited participants to the National Research Council Workshop.

Panel Format

The program will consist of an hour and 15-minute panel discussion organized as follows:

1. Overview of Simulator-Induced Sickness (L. Frank)

   ° Nature of the Problem
   ° Symptoms
   ° Incidence
   ° Common Characteristics across studies
   ° Reasons for concern

2. Etiology (R. Kennedy)

   ° Stimulus characteristics
   ° Response characteristics (behavioral/physiological)
   ° Anatomical structures of significance
   ° Predisposing factors
   ° Theories of motion sickness

3. Simulator Design and Use Factors Contributing to Simulator-Induced Sickness (G. Ricard)

   ° Cue characteristics
     e.g., fidelity, delays
   ° Inertial System Characteristics
     e.g., resonant frequencies, washout
   ° Visual System Characteristics
     e.g., field-of-view, vection

4. Recommendations of National Research Council's Workshop on Simulator Sickness (R. Hennessy, M. McCauley)

   ° Countermeasures
   ° Research
   ° Other

5. Summary, Synthesis, and Perspective (H. Leibowitz)

Affective Determinants of Job Perceptions

Kurt Kraiger
University of Colorado at Denver

Abstract

The Job Characteristics Model of Hackman and Oldham (1976) has served as a useful guide for designing jobs to be more motivating and satisfying. It is argued, however, that the job perceptions of incumbents may be biased or influenced by internal affective states or moods. This assertion is supported by the results of two studies. The first was a laboratory study in which job perceptions were more favorable when subjects were artificially placed in a good mood. In the second study, the overall job perceptions of a sample of city government workers were found to be predictable from both their job satisfaction and mood states.

Introduction

During the past fifteen years, there has been considerable research interest in the job perceptions and job attitudes of workers. Much of this interest was stimulated by the Job Characteristics Model of Hackman and Oldham (1976). The Job Characteristics Model makes three primary assertions: 1) All jobs may be accurately characterized or described by a small number of attributes such as significance or autonomy; 2) the degree to which any one job possesses these qualities determines the intrinsic interest or motivating potential of that job; and 3) the higher a job's motivating potential, the higher will be the satisfaction, motivation, and performance of its incumbents. Subsequent research has offered at least moderate support for these assertions (see Roberts & Glick, 1981, for a review).

A question of current concern is whether job perceptions of workers are totally objective or whether these measures are influenced by other factors, including attitudes or dispositions of the workers. Employees' internal states may be a potential determinant of perceived job characteristics. Hanser and Muchinsky (1978) have suggested that as external information sources (i.e., the work environment) become more ambiguous, employees should be more influenced by internal attitudinal or affective informational sources when forming job perceptions. Cognitive consistency theory posits that individuals seek to maintain balance between their attitudes, beliefs, and behavior. Consequently, workers may align environmental (job) perceptions with their prevalent job attitudes (James & Jones, 1980). A number of recent studies have shown that satisfied workers describe their jobs in more positive terms than do dissatisfied workers (e.g., Caldwell & O'Reilly, 1982; James & Jones, 1980).

Workers' moods or affective states are suggested here as another possible influence on job perceptions. Affective states are defined as relatively mild emotion-like states which may alter subsequent non-affective thoughts or behaviors. (Examples of affective states are pleasantness or irritation.) Affect may influence the observation, storage, recall, or evaluation of job-related information in at least three ways (Schwarz & Clore, 1983). First, the affect itself may serve as in information source about the job; workers perceive their job positively because they feel good while they are there. Secondly, moods may direct the individuals' attention to environmental

conditions which are plausible causes of their good feelings. Thirdly, affective states may serve to increase the availability of mood-congruent information in memory (Isen, Shalker, Clark, & Karp, 1978); workers in a good mood primarily remember only positive job events.

In recent years, considerable research has demonstrated that mood states influence subsequent thoughts, perceptions, and behavior. While in a good mood, subjects have been found to process information differently than control subjects on a decision-making or problem-solving task (Isen & Means, 1983) and recall more positive words than negative words in a free recall task (Isen et al., 1978). More germane to the present discussion, Isen and Shalker (1982) found subjects in a positive mood rated slide pictures of familiar scenes more favorably than did subjects in a neutral or negative state. Notably, Isen and Shalker demonstrated that positive mood states did not debilitate subjects' ability to discriminate among slides but tended to raise the overall level of their ratings.

Given the prior research on other biases in job perceptions, it is hypothesized that perceived job characteristics are influenced by incumbents' affective states. To test this hypothesis, a laboratory study was conducted in which positive and neutral affect subjects were given both an interesting and a dull task to perform. The important research questions were: 1) Would both positive and neutral affect subjects distinguish between an interesting and a dull task? and 2) regardless of task, would positive affect subjects characterize the tasks in more favorable terms than would neutral affect subjects?

Given that an effect could be produced and detected under artificial conditions, the next logical step would be to demonstrate its existence in a true field setting. This was attempted using survey data obtained from a sample of municipal workers. With this database, additional research questions regarding the affect construct were examined: 3) Could individual differences in chronic or long-term mood states be reliably and validly measured?; 4) if so, to what extent are these chronic states distinct from momentary, fleeting moods?; and 5) what are the relative contributions of job satisfaction, momentary mood states, and chronic mood states to the prediction of job perceptions?

Study 1

Study 1 was a laboratory study conducted on 80 student volunteers. Subjects were recruited for an experiment on the selection of graduate teaching assistants for a large midwestern university.

All subjects performed both an enriched or interesting task and an unenriched or boring task. (Task order was counterbalanced.) In the enriched task, subjects were given no explicit rules or procedures, were allowed to make hiring decisions, and could work at their own pace. In the unenriched task, subjects received restrictive directions, simply coded candidate information, and were paced by an audiotape. Each task lasted 20 minutes. Measures of task perceptions were collected after each task.

Task design was crossed with affective state. Control subjects received no affect induction and were considered "neutral" in mood. Positive affect subjects were shown a four-minute videotape of TV and movie bloopers before each task. The tapes were introduced as pilot material for a future study and were confirmed as a successful method of affect induction in an earlier study.

The results showed that task design did have a strong effect on perceived task characteristics. Mean ratings were significantly higher in the enriched

condition for all five task dimensions (task significance, task feedback, task identity, autonomy, and skill variety) and for the overall task summary score, Motivating Potential Score (MPS).

Of greater interest are the results of the effects of affective state on task perceptions. Regardless of task design, positive affect subjects gave significantly higher ratings of task characteristics than did neutral affect subjects ($F=6.73$; $df=1,76$; $p<.01$ for MPS ratings). Effects due to mood were strongest for task significance, task feedback, and skill variety. These results support one important hypothesis: Given identical jobs or tasks, workers in a good mood will describe that job more favorably than neutral affect workers.

Study 2

Study 2 was conducted on a heterogeneous sample of workers from a large city government. Seventy-four individuals returned their surveys and were included in the analyses (response rate = 55%). Subjects ranged in age from 21 to 67 years old. Females constitued 62% of the sample. Organizational status ranged from custodial workers to executives within the mayor's cabinet.

Data were collected on the following variables: Momentary mood while completing the survey (measured by three bipolar scales), Motivating Potential Score (MPS; summed over ratings of the five core dimensions), overall job satisfaction, and chronic mood. Chronic mood was measured in two ways. First, respondents were given behavioral descriptions of persons who were always in a good mood or always in a bad mood. Subjects then listed the names of ten co-workers and indicated whether they themselves were more or less often in a good mood then each peer. The self-rating of chronic mood was determined by the number of co-workers that they rated themselves higher than (a similar procedure for measuring need for achievement was developed and validated by Holmes and Tyler, 1968). Additionally, subjects rated each co-worker on a five-point scale, indicating the extent to which that person was typically in a good mood. Both the indirect self-rating of chronic mood state and the average peer rating were more highly correlated with each other ($r=.57$) than either was with the direct momentary mood rating ($r=-.08$ for self-rating, $r=.26$ for peer rating). Along with the high level of agreement among peers rating the same worker, this evidence suggests that individual differences in chronic mood states can be reliably and accurately measured and that these long-term states are at least partially independent of momentary states.

The important question dealt with the relative contribution of the various affective variables in the prediction of job characteristics. This was examined using hierarchical multiple regression analysis to predict MPS scores. Table 1 shows the increments in variance accounted for in MPS by adding several different combinations of predictors. In all cases, three individual difference variables (age, division, and rank) were entered first as a control for extraneous variance. It should be noticed that the results were similar using either chronic mood measure. Results are presented for peer ratings.

As can be seen from Table 1, all three affective variables were predictive of MPS ratings. Consistent with prior research, knowledge of worker job satisfaction significantly and greatly increased the ability to predict job perceptions (analysis A or C). Mood states at the time in which perceptions were measured (momentary mood) were also a significant and strong predictor of job perceptions, even when job satisfaction was already accounted for (analysis A and B). Results for chronic mood states were less impressive. Chronic mood states alone did contribute significantly to the prediction of MPS scores beyond

Table 1

Relative Contribution of Job Satisfaction, Momentary and Chronic Mood States to the Prediction of MPS Scores

Analysis	Step	R	R^2	ΔR^2	$F(\Delta R^2)$
	1. Age, Sex, Rank	.23	.05	.05	1.18
A.	2. Job Satisfaction	.54	.29	.24	21.52***
	3. Momentary Mood	.62	.38	.14	9.95***
	4. Chronic Mood	.63	.39	.01	1.17
B.	2. Momentary Mood	.48	.23	.18	15.74***
	3. Chronic Mood	.51	.26	.03	2.44*
C.	2. Job Satisfaction	.54	.29	.24	21.52***
	3. Chronic Mood	.56	.31	.02	2.56*
D.	2. Chronic Mood	.35	.12	.07	5.32**

N = 70   * $p < .05$ ** $p < .01$ *** $p < .001$

that accounted for by individual difference variables (analysis D). However, when either job satisfaction or momentary mood states were also controlled for, chronic states accounted for only a neglible amount of additional variance (analyses A, B, and C).

Discussion

The Job Characteristics Model remains an important contribution to organizational psychology. It provides both a method for diagnosing jobs and a method for changing those jobs in ways which should improve the satisfaction, motivation, and performance of incumbents. The present paper cautions that leaders should be careful not to weigh perceptions of job attributes too heavily when examining morale or performance problems in their units. Low job perceptions may be the result of a morale problem rather than the cause of it. Ratings of job characteristics are likely to be influenced by several attitudinal and affective dispositions of incumbents. Individuals who feel good are likely to respond to their jobs in a more positive manner than individuals who feel bad.

In this era of increased emphasis on human relations skills in supervision and admiration for the "Japanese management style," it should not be surprising to discover that workers' feelings do make a difference. Chronic mood states can be identified, but do not predict job perceptions. This would suggest that solutions to morale or performance problems do not lie in identifying and selecting only "happy people;" rather, emphasis should be placed on improving the day to day feelings of all unit members. Some suggestions for improving affective states might include administering noncontingent rewards (e.g., compliments from supervisors), building associations in workers' minds between the primary work setting and positive experiences, and encouraging the worker to mentally dwell on previous life or work events. Each of these methods have proven to be effective methods of increasing positive affect under experimental conditions. Their value for improving job attitudes in real world settings has yet to be established. While most psychologists would agree that attitudes are primarily affective in nature, it is paradoxical to note that all methods

of attitude change are cognitively or behaviorally-based. There is no widely accepted theory of affectively-based attitude change.

References

Caldwell, D.R., & O'Reilly, C.A. (1982). Task perceptions and job satisfaction: A question of causality. Journal of Applied Psychology, 67, 361-369.

Hackman, J.R., & Oldham, G.R. (1976). Motivation through the design of work: Test of a theory. Organizational Behavior and Human Performance, 16, 250-279.

Hanser, L.M., & Muchinsky, P.M. (1978). Work as an information environment. Organizational Behavior and Human Performance, 21, 47-60.

Holmes, D.S., & Tyler, J.D. (1968). Direct versus projective measurement of achievement motivation. Journal of Consulting and Clinical Psychology, 33, 712-717.

Isen, A.M., & Means, B. (in press). The influence of positive affect on decision-making strategy. Social Cognition.

Isen, A.M., & Shalker, R.E. (1982). The effect of feeling state on evaluation of positive, neutral, and negative stimuli: When you "accentuate the positive," do you "eliminate the negative"? Social Psychology Quarterly, 45, 58-63.

Isen, A.M., Shalker, T.E., Clark, M.S., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? Journal of Personality and Social Psychology, 36, 1-12.

James, L.R., & Jones, A.P. (1980). Perceived job characteristics and job satisfaction: An examination of recipricol causation. Personnel Psychology, 33, 97-135.

Roberts, K.H., & Glick, W. (1981). The job characteristics approach to task design: A critical review. Journal of Applied Psychology, 66, 193-217.

Schwarz, N., & Clore, G.L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. Journal of Personality and Social Psychology, 46, 513-523.

Expectations, Job Perceptions and Disconfirmation Among Air Force Academy Graduates

Sharon C. Conley

Abstract

This study examined the expectations of Air Force Academy cadets, their later job perceptions as officers, and the discrepancy between these expectations and perceptions (disconfirmation). Using Hackman and Oldham's job characteristics model as a conceptual base, this research examined expectations about five job characteristics (autonomy, feedback, skill variety, task significance, and task identity). Subjects were 110 male and female 1980 graduates of the Air Force Academy (the first class that graduated women). Data on expectations were collected before the subjects' graduation from the Academy: two years later, job perceptions were assessed by means of mailed questionnaires. Results showed that job perceptions differed from expectations only with respect to autonomy--which was higher than expected.

Since previous research suggests that males and females--and those in flying versus support positions--may differ in their expectations and job perceptions, these sub-groups were also examined.

Introduction

The Air Force Academy, as a primary source of commissioned officers, aims to provide cadets with the necessary skills and background experiences to serve as career officers. The Academy experience, as a socializing influence, also provides cadets with certain expectations concerning their future Air Force jobs. Organizational socialization theory suggests that these expectations are then "tested" against organizational reality once these individuals enter the job setting. Previous research suggests that if there is a discrepancy between expectations and later job perceptions (disconfirmation), dissatisfaction can occur. Specifically, Wanous (1980) suggests that traditional organizational recruiting practices typically lead to inflated expectations--and reduced satisfaction--among newcomers to an organization.

Therefore, it appears important to examine the job expectations and later job perceptions of cadets to determine the extent to which disconfirmation is present, and also to determine the relationships between expectations, disconfirmation and work outcomes (e.g., satisfaction and motivation).

This study will focus on expectations concerning the characteristics of the work that is performed, as job characteristics are believed to be a major factor in employee motivation and satisfaction. Using Hackman and Oldham's (1976) job characteristics

model as a conceptual base, this study will examine five job characteristics: autonomy, skill variety, task significance, feedback and task identity.

Finally, previous research suggests that job characteristics and work outcomes may be different for male and female Air Force officers (DeFleur, 1981) and for officers in flying versus support functions (Wood, 1980). Therefore, this research will also examine expectations, job perceptions, disconfirmation, and work outcomes for these subgroups.

Method

Subjects were 53 male and 57 female Air Force officers (lieutenants) who graduated from the Air Force Academy in 1980. The subjects were from an original sample of approximately 300 cadets who were selected by Academy researchers for a study concerning the integration of the first class of women into the Academy.

Data concerning expectations were collected from 175 cadets one month prior to their graduation from the Academy. Two years later, survey instruments measuring perceptions of job characteristics and work outcomes were mailed to the participants of the first data collection. One hundred and ten officers responded, yielding a response rate of 63%.

The Job Diagnostic Survey (JDS), developed by Hackman and Oldham in 1975, was utilized to measure perceptions and expectations about the following five job characteristics: skill variety (whether the job requires a variety of different activities in carrying out the work); task identity (whether the job requires following a task through, from beginning to end); task significance (whether the job has a substantial impact on others); autonomy (whether the job provides freedom and discretion to the individual in carrying out the work); and feedback (whether the job itself provides the worker with information about the worker's performance). Since the JDS scales measure perceptions about the above characteristics, one item from each scale was reworded to measure expectations.

Disconfirmation of job characteristics was measured by subtracting each subject's expectations score from the corresponding job perceptions score. In addition, in the second phase of the study, subjects were asked to indicate the degree to which they felt their expectations concerning each job characteristic were met. Thus, two measures of disconfirmation were utilized in the study.

Finally, three work outcomes were assessed--job satisfaction, motivation and job involvement. Satisfaction and motivation were assessed with the JDS. Lodahl and Kejner's (1965) job involvement scale was utilized to measure job involvement (personal involvement in one's work).

Results

Means on expectations range from 4.1 (autonomy) to 5.2 (task significance and feedback). Mean scores for job perceptions range from 4.7 (autonomy) to 5.2 (task significance). All variables were measured on 7-point scales. The means indicate that the officers, on the average, have a substantial amount of the job characteristics in their jobs.

T-tests were utilized to compare the means on expectations with the corresponding perceptions of job characteristics. The only significant difference is on autonomy ($t=4.12$, $p<.001$). The results indicate that perceptions of autonomy are significantly higher than expectations.

Pearson product-moment correlation coefficients were utilized to examine the relationships of expectations--and disconfirmation of expectations--with the work outcomes. Expectations about task significance are positively related to satisfaction (.21) and job involvement (.20); and expectations about feedback are positively related to satisfaction (.28). In addition, the two disconfirmation measures show some consistency in the results, in that both measures of disconfirmed skill variety are positively related to motivation, satisfaction, and job involvement. Therefore, perceiving more skill variety in the job than expected is associated with higher motivation, involvement, and satisfaction--perceiving less variety than expected is associated with less positive work outcomes. Results concerning disconfirmation about the other four job characteristics are not as consistently strong.

Males versus Females

Males and females were not found to be significantly different on expectations, job perceptions, or disconfirmation. However, males are higher on job involvement than are females ($t=2.46$, $p<.05$).

Flying versus Support Officers

Flyers perceive more skill variety and feedback--but less autonomy--in their jobs than do support officers ($t=2.08$, $p<.05$; $t=3.44$, $p<.01$; $t=^{-}3.05$, $p<.05$, respectively). Flying officers are also higher on job involvement than are support officers ($t=5.72$, $p<.001$). No significant differences were found on expectations or disconfirmation.

Discussion

This study examined the expectations, disconfirmation of expectations, and job perceptions of Air Force Academy graduates.

The findings revealed that, except for autonomy, the officers' perceptions of job characteristics were not significantly different from their expectations. These results are not consistent with Wanous's

(1980) proposition that newcomers enter their jobs with inflated expectations. The findings may be attributable to the realistic information cadets receive at the Academy. It is likely that Academy instructors are a source of realistic information, since they are all Air Force officers. In addition, the Air Force work experience that many cadets receive during their summer months may provide a "realistic preview" of their future jobs.

It is possible that officers perceive more autonomy than they expected because they may base their expectations on their Academy experience. Because there is limited autonomy at the Academy, cadets may tend to underestimate the autonomy they will receive in their jobs.

Expected task significance and feedback were found to be related to satisfaction, and expected task significance was also related to job involvement. Moreover, the findings concerning disconfirmation suggest that officers who do not experience as much skill variety as they expected are less likely to feel satisfied, motivated and involved in their work. Since it is particularly important that expectations concerning skill variety be met, attempts should be made to inform newcomers about this facet of their jobs.

The comparison of males and females showed that although women are not less satisfied with their jobs, they do have lower job involvement, that is, they are less personally involved in their work. As cadets, the women anticipated problems with combining career and family roles (DeFleur, 1981), which could be contributing to lower job involvement. Since DeFleur (1981) also found that women cadets objected to the exclusion of women from combat roles, dissatisfaction with Air Force policies could also be contributing to their lower job involvement.

The finding that flyers have more skill variety and feedback, but less autonomy than do support officers is not surprising. Learning to fly involves mastery of a complex task, which demands numerous skills and abilities. In addition, the plane's performance provides immediate feedback to the pilot. Wood (1980) suggests that pilots have less autonomy than do support officers because the mistakes of support officers are less likely to be disastrous. Finally, with regard to work outcomes, the finding that flyers have higher job involvement may be because flying is seen as the core activity of the Air Force. Therefore, the flying officer may be more likely to perceive his or her job as significant to the Air Force mission.

References

DeFleur, L.B. Report submitted to the Air Force Office of Scientific Research, 1981.

Hackman, J.R., and Oldham, G.R. Development of the job diagnostic survey. Journal of Applied Psychology, 1975, 60, 159-170.

Hackman, J.R., and Oldham, G.R. Motivation through the design of work: test of a theory. Organizational Behavior and Human Performance, 1976, 16, 250-279.

Lodahl, T.M., and Kejner, M. The definition and measurement of job involvement. Journal of Applied Psychology, 1965, 49, 24-33.

Wanous, J.P. Organizational entry. Reading: Addison-Wesley, 1980.

Wood, F.R. Air Force junior officers: changing prestige and civilianization. Armed Forces and Society, 1980, 6, 483-506.

The author wishes to acknowledge that the expectations data were collected under the direction of Maj. Frank R. Wood at the Air Force Academy. Maj. Wood and other Academy staff in the Department of Behavioral Sciences and Leadership also provided planning and technical assistance for the second data collection phase.

The Effects of REFORGER Exercise Participation On Soldiers' Attitudes

Dennis J. Grill, PhD
William Beaumont Army Medical Center

Abstract

Thousands of DOD personnel annually participate in REFORGER (REturn of FORces to GERmany) exercises along with other NATO Allies. Pre and post REFORGER "Unit Status Questionnaires," similar to surveys utilized by Israeli Defense Force (IDF) battle psychologists, were administered to two hundred and fifty two soldiers in an armored cavalry squadron which participated in REFORGER '83, Exercise "Confident Enterprise." Analysis focused on determining the effects of the unit's participation in REFORGER. A factor analysis of both the pre and post REFORGER "Unit Status Questionnaires" yielded a set of five factors: (1) View of the Unit, (2) Unit Communication, (3) Self-Evaluation, (4) Evaluation of the Enemy, and (5) Individual Morale. Participation in REFORGER had no significant effect on any of these factors. However, further analysis revealed that there were significant differences between certain groups on a number of dimensions such as (1) previous REFORGER participation, (2) type of MOS, (3) pay grade, and (4) unit. Implications for unit training and for the utilization of mental health personnel in combat are examined.

Introduction

It has long been assumed that a unit's participation in a field training exercise, such as REFORGER, has a positive effect on that unit's morale, confidence in its leaders and equipment, and overall combat readiness. After all, deploying to the Federal Republic of Germany, drawing combat equipment, convoying by rail and road, participating in such a large scale tactical exercise, turning-in equipment and returning to the United States, in effect, affords units the opportunity to practice one of their major wartime missions. But is this assumption accurate, or is it more of a wish on the part of exercise planners and commanders at various levels.

Israeli Defense Force (IDF) battle psychologists have effectively utilized unit survey data to provide commanders with feedback about their units during the 1973 Yom Kipper War (Greenbaum, Rogovsky & Shalit, 1977) and the 1982 War in Lebanon (Belenky, Tyner & Sodetz, 1983). The present

study is an effort to conduct similar "action research" which would be value and relevance to commanders and at the same time, examine the effects of REFORGER participation on soldiers' attitudes about morale. In addition, soldiers' attitudes about morale were examined in terms of other relevant variables. Does previous REFORGER experience with the same unit, the type of job (MOS) a soldier performs in the unit, his pay grade, marital status or individual unit, have any effect on morale?

Method

The survey instrument, entitled "Unit Status Questionnaire," included a series of seventeen questions to which each soldier responded on a Lickert scale format from 1 to 5, 1 being the positive end of the continuum, and 5 the low. Thirteen of the questions were taken from the Israeli morale survey questionnaires as reported by Miller (1983). These thirteen questions

TABLE 1
DESCRIPTIVE DATA
UNIT STATUS QUESTIONNAIRE

1. Unit		3. Participated in REFORGER '82	
A	58	Yes	51
B	46	No	201
C	29		252
D	34		
HOW	29	4. Military Occupational Specialty	
HHT	56	Combat MOSs	190
	252	Support MOSs	62
			252
2. Pay Grade			
E1-E4	163	5. Marital Status	
E5-E7	82	Single	111
E8/Officer	7	Married	126
	252	Divorced/Separated	15
			252

dealt with (1) morale, individual and unit, (2) unit readiness for combat, (3) level of trust in unit leaders, in the unit as a whole, in friends, in self and in equipment, (4) self-evaluation of soldiering ability, (5) fear of combat and (6) knowledge of unit mission. Four additional questions which tapped unit cohesion were taken from Manning(1980). The basis for selection of the individual questions was their perceived relevance to commanders in the field, preparing for combat. Soldiers were also asked to specify their unit/platoon/section, rank, military occupational specialty (MOS), marital status, and whether or not they had been on REFORGER with the same unit the previous year. Table 1 contains a summary of this

descriptive/demographic data.

The unit selected to participate in the study was an armored cavalry squadron stationed at Fort Bliss, Texas. The Pre-REFORGER Unit Status Questionnaires were administered approximately ten days prior to deployment. The Post-REFORGER Unit Status Questionnaires were completed within two weeks after the unit's return to the United States. Two hundred and fifty two matched (pre and post REFORGER) Unit Status Questionnaires were collected.

Results

A factor analysis of both the pre and post REFORGER questionnaires yielded the same five factors: (1) view of the unit, (2) unit communica-

TABLE 2
FACTOR ANALYSIS RESULTS
PRE/POST LOADINGS

Factors	Pre-REFORGER Loading	Post-REFORGER Loading
Factor 1. View of the Unit		
-Trust in the Unit	0.837	0.833
-Unit Combat Readiness	0.778	0.827
-Unit Combat Confidence	0.768	0.844
-Trust in Unit Leaders	0.731	0.821
-Unit Morale	0.607	0.765
-Trust in Equipment	0.597	0.687
-Combat Preparatory Talk	0.534	0.530
-Trust in Unit Friends	0.534	0.485
-(Individual Morale)	(-----)	(0.629)
Factor 2. Unit Communication		
-Communication w/Platoon Leader	0.722	0.767
-Communication w/Squad Leader	0.660	0.608
-Communication w/Unit Commander	0.603	0.592
-Knowledge of Unit Mission	0.447	0.518
-(Individual Morale)	(0.481)	(-----)
Factor 3. Evaluation of Self		
-Trust in Self	0.829	0.809
-Self-Evaluation as a Soldier	0.811	0.805
Factor 4. Evaluation of the Enemy		
-Evaluation of the Soviet Soldier	0.688	0.699
-Fear of Combat	0.481	0.631
Factor 5. Individual Morale	0.481	0.629

tion, (3) evaluation of self, (4) evaluation of the enemy and (5) individual morale. Table 2 depicts the highest factor loading for each individual question on a factor analysis of pre and post questionnaire data. Examination of the loadings reveals the stability and strength of the clustering of all the questions except one, individual morale. This question shifted from Factor 2 (Unit Communication) in the pre-REFORGER analysis to Factor 1 (View of the Unit) in the post-REFORGER analysis; individual morale also gain in strength (0.481 to 0.629). For these reasons, individual morale was examined as a separate factor.

TABLE 3
REPEATED MEASURES
ANALYSIS OF VARIANCE
SIGNIFICANT VALUES OF F (p .05)

Source	Factor 1 View of Unit	Factor 2 Unit Commo	Factor 3 Self Eval	Factor 4 Enemy Eval	Factor 5 Ind Morale
Pre/Post REFORGER On All Variables On All Factors	-----	-----	-----	-----	-----
REFORGER '82 Participation	-----	4.52	-----	4.98	-----
MOS(Cbt/Spt)	32.79	6.39	-----	-----	15.67
Pay Grade	9.60	8.47	-----	-----	3.87
Marital Status	-----	-----	-----	-----	-----
Unit	19.18	3.86	-----	2.49	3.76

A repeated measures analysis of variance was performed on the pre and post REFORGER data over all five factors and over five demographic variables. These five demographic variables were: (1) REFORGER '82 participation (yes, no), (2) military occupational speciality (MOS) (combat, support), (3) pay grade (E1-E4, E5-E7, E8/Officer), (4) marital status (single, married, divorced/separated), and (5) unit (A, B, C, D, HOW, HHT). The results of this analysis are listed in Table 3.

There were no significant differences between pre and post measures on any of the five factors or over any of the five demographic variables. Participation in REFORGER had no effect on morale. In several ways, though, this finding is not surprising. The tactical exercise phase of REFORGER is, in reality, a corps level exercise, not one geared to the individual soldier. Following the tactical exercise phase, three to four weeks are spent cleaning and turning-in equipment. This turn-in process usually produces frustration rather than increasing morale or esprit-de-corps. Any increase in morale gained from the tactical exercise phase is soon lost during the turn-in of equipment. Additionally, during turn-in soldiers have only one question on their minds -- "When can I go home?"

While there were no significant differences between pre and post REFORGER measures, there were several significant differences among the demographic variables on various factors. Those soldiers who participated in REFORGER '82 with the same unit felt that unit communication between them and the commander, platoon leader and squad leader was much better than those who did not (Factor 2). Interestingly also, these same soldiers perceived an enemy as a more competent adversary than those who had not been on REFORGER '82 (Factor 4). Soldiers in combat MOSs (11-19 career management fields) viewed their unit more positively (Factor 1), felt that unit communication was better (Factor 2) and their individual morale (Factor 5) was higher than those soldiers with support MOSs (all other career management fields). Commanders may need to focus more of their attention on soldiers filling support roles. Another significant variable was pay grade. E8/Officers viewed their units more positively (Factor 1), felt that unit communication was better (Factor 2) and their individual morale was higher (Factor 5) than junior enlisted and non-commissioned officer personnel. These results may suggest that first sergeants and officers are out of touch with the junior and senior enlisted ranks, or at least are viewing their unit through rose-colored glasses. Marital status had no effect on any of the five factors.

The final variable in which significant differences were found, and probably the most important, was unit. Units differed on four of the five factors. What is important here is not how the company-sized units of this particular armored cavalry squadron differed, but rather the fact that they did differ. The squadron commander can utilize such unit information to aid in planning for unit training or selecting a unit for a specific mission. Mental health personnel, serving as consultants to command, utilizing similar survey techniques, can help commanders maximize the potential of their units by providing relevant feedback about soldier attitudes and morale.

References

Belenky, G. L., Tyner, C. F., & Sodetz, F. J. Israeli battle schock casualties: 1973 and 1982. Washington, D. C.: Walter Reed Army Institute of Research, 1983.

Greenbaum, C. W., Rogovsky, I., & Shalit, B. The military psychologist during wartime: A model based on action research and crisis intervention. Journal of Applied Behavioral Science, 1977, 13, 7-21.

Manning, F. J. Assessment of unit cohesion and its relation to unit performance. Paper presented to the Anglo-American Psychiatry Symposium, London, England, 1980.

Miller, J. A. OCONUS trip report - Attendance at the third international conference on psychological stress in time of war and peace. Personal Communication, 1983.

Changes in Attitudes Toward Women at the Air Force Academy

Alan C. Bridges
United States Air Force Academy

Abstract

Sex-role attitudes were measured at the Air Force Academy by Marshak, DeFleur, and Gillman (1976), using the Attitudes toward Women Scale (Spence & Helmreich, 1972). The attitudes were significantly less egalitarian than those held by male college students at the University of Texas at Austin. Festinger's Cognitive Dissonance Theory (1957) predicts that as perceptions of mandated social change decrease, attitudes held by cadets would become more egalitarian. In cadets sampled in 1983, males are significantly more egalitarian than in 1976, but there's no significant difference between 1983 cadets and 1980 college students. There's a significant trend in male cadets: the longer he's been at the Academy, the more egalitarian he is. As expected, comparison between 1983 male and female cadets showed women were significantly more egalitarian than men.

Introduction

Background

Since the integration of women into the service academies in 1976, several studies have tested cadets' sex-role attitudes. These studies came to the same conclusion: the men's attitudes toward women are significantly more conservative than those held by their civilian counterparts.

The first study (Vitters & Kinzer, 1976) was done at the U.S. Military Academy. It found that the all-male upperclass was significantly more conservative than civilians in the attitudes they held. The second study, by Marshak, DeFleur and Gillman (1976), found the same conservative trend in the Air Force Academy cadets. A unique study was done at the Coast Guard Academy by Rottman (1980). She attempted to change male cadet attitudes through cognitive dissonance. The results weren't significant, unfortunately, but the study proved cadets at the Coast Guard Academy still held conservative attitudes as late as 1980.

Festinger's Cognitive Dissonance theory (1957) predicts that people behave to reduce dissonance created by conflicting attitudes or beliefs. During mandated social change, however, there's less pressure to reduce

dissonance created by conflicting attitudes or beliefs. In the context of a service academy, accepting women was a mandated social change, so there's little pressure to reduce any dissonance.

Over time, the perceptions of mandated social change decrease, since women aren't a recent addition any more. Marshak, et al (1976), predicted that the acceptance of women in the service academies would be similar to the earlier integration of blacks: early resistance and later acceptance.

In 1978, Heilman and Guzzo reported that perceptions of the cause of work success could act as a mediator of sex-role biases. Basically, they found that if women were perceived as being successful due to their ability, they were perceived as being successful as men at the same task. At the AFA in recent years, women have held increasingly responsible jobs in the cadet wing. Requirements are stringent: jobholders must have good grades and a good military bearing. Military performance ratings might be written off as subjective but grades are an accepted measure of performance. In 1981-82 one of the wing commanders was a woman. With such successes by women, which have to be attributed to their ability, one could expect attitudes to become more egalitarian.

Today, with every class at the Academy only being in contact with integrated classes, women aren't a new phenomenon in the cadet wing. There's little media coverage on the subject of forced integration, and the Academy is doing nothing to publicize past troubles. Dissonance theory would predict that attitudes would become more egalitarian as perceptions of mandated social change decrease.

Method

Subjects in this study were 120 male cadets (30 from each class) and 30 female (5 seniors, 10 juniors, 9 sophomores, 6 freshmen. They were tested during a core course to get a random sample. The testing apparatus was the 25-item version of the Attitudes toward Women Scale (AWS) (Spence & Helmreich, 1972) used in the 1976 AFA studies. A score for the 15-item AWS was extracted for comparison with 1980 civilian data (Spence, Personal Communication).

Using a one-way analysis of variance (ANOVA), the male cadet sample was tested for differences between the classes. A separate one-way ANOVA was done on male-female cadet attitude differences. Comparisons were made between the 1983 male cadet sample and both the 1976 cadets and 1980 civilian students using two t-tests. The 1983 scores were taken for all four classes at the Academy, while the 1976 scores were taken from seniors, juniors, and freshmen only. The standard deviation from the 1983 cadet sample was used in the comparison with 1980 college students because civilian dispersion was not available. The α-level for each t-test was held at .02 to get an overall α-level of .05. The statistics were computed by hand or by using the Statistical Package for the Social Sciences (SPSS) (Nie, et al, 1975).

Results

The results of the first ANOVA are interesting: the longer a male cadet has been at the Academy, the more egalitarian he is as measured by AWS. The difference between the classes was significant ($F=3.906$, $df=3, 116$, $p< .05$). It's encouraging to find that the Academy's curriculum has beneficial side effects.

As expected, the women at the Academy are more egalitarian than the men. The difference was significant ($F=27.5$, $df=1, 148$, $p< .001$). To want to enter a predominantly male service academy, a woman has to know she can do as well as a man.

1983 male cadets are significantly more egalitarian than their 1976 counterparts ($t=3.69$, $df=119$, $p< .001$). According to data from Dr. Janet Spence, there wasn't a comparative trend between 1976 and 1980 college students: the biggest change was between 1972 and 1976 students. One possible reason this change was slow to show up in cadets is the Academy's selection process. The Academy likely attracts people with more conservative values, so it took longer for the change to appear.

The most encouraging finding was that 1983 male cadets aren't significantly different from 1980 college students if males haven't shifted ($t=1.145$, $df=119$, N.S.). This is exciting because of what it says about society and the changes at the Air Force Academy: society is more egalitarian, and the Academy is changing, too.

Discussion

The results bear out the prediction of Festinger's Cognitive Dissonance theory, which is that as perceptions of mandated social change decrease, attitudes held by male cadets would tend to become more egalitarian. There may be several reasons for this egalitarian shift. The men may see the women's favoritism. Over time, the perceptions of forced compliance have decreased, so there's more pressure to reduce dissonance. The possibility that society has become more egalitarian was refuted by the survey by Helmreich, Spence, and Gibson (1982). They found that society leveled off over the time the cadet comparison was made.

A more plausible explanation for the delay in the egalitarian shift is the Academy's selection process. Service academies usually attract students who feel traditional and probably hold conservative ideas about their country and, consequently, women also. As society changes, though, it's good to see the Air Force Academy is keeping pace.

That women at the Academy believe women are more capable than the men at the Academy is to be expected. Being a woman in a traditionally male role implies a belief in women's abilities. The differences between the male classes is encouraging. Possibly the men enter with traditional views and have them refuted by seeing women excel in an extremely tough environment.

One obvious question jumps out after these findings: what are the attitudes at the other service academies? If they're less egalitarian, is it due to the selection process or something the Air Force Academy is doing that they aren't? If they're more egalitarian, is it due to the selection process, again, or to something the other academies are doing that the Air Force Academy isn't? This is a promising area for further research.

Acknowledgements

I'd like to thank Robert A. Gregory, Major, USAF, for his comments on the drafts. I'd especially like to thank William P. Marshak, Captain, USAF, for his help with the research and with revising several drafts.

References

Baron, R.A., & Byrne, D. Exploring Social Psychology. Boston: Allyn and Bacon, Inc., 1982.

Heilman, M.E., & Guzzo, R.A. The perceived cause of work success as a mediator of sex discrimination in organizations. Organizational Behavior and Human Performance, 1978, 21, 346-357.

Helmreich, R.L., Spence, J.T., & Gibson, R.H. Sex-role attitudes: 1972-1980. Personality and Social Psychology Bulletin, 1982, 8, 656-663.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. Applied Statistics for the Behavioral Sciences. Boston: Houghton Mifflin Co., 1979.

Marshak, W.P. Integration of women into the United States Air Force Academy: attitudes during the initial year. Unpublished masters project, University of North Dakota, 1978.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. Statistical Package for the Social Sciences. New York: McGraw-Hill Book Co., 1975.

Rottman, M.N. Effects of a self-administered feminine literature package on attitudes toward sex-role stereotypes among third-class cadets at the United States Coast Guard Academy. Unpublished doctoral dissertation, University of Connecticut, 1980.

AD-P003 262

Risk-Taking Performance of Military Personnel: Sex Differences and Practice Effects

Gerald A. Hudgens, Ph.D.
Linda Torsani Fatkin
US Army Human Engineering Laboratory
Aberdeen Proving Ground, Maryland 21005

An initial search of literature dealing with sex differences and performance variables led to several conclusions about past research in the area: 1) a large number of studies have been reported (Hudgens & Fatkin, 1980, cited about 1500 references); 2) most of those studies were conducted by researchers whose primary interest was the performance variable rather than the sex variable which appeared most often to be included only incidentally; and 3) most of the studies represented one-time efforts both in that repeated measures were not obtained within the investigations and in that the investigations were not followed up to determine the reliability of, generality of or factors responsible for the sex differences observed. The present program was established to perform more intensive investigations of sex differences on performance variables of potential importance to military applications.

The literature search also led to the identification of "risk-taking" behavior as one performance variable which could likely yield sex differences of potential importance within military contexts. Our interest in this area was sparked mainly by reports on driving behavior which have consistently demonstrated a more conservative attitude toward risk taking in women drivers (e.g., Ebbesen & Haney, 1973). Our initial efforts have involved attempts to test for sex differences in simplified computer simulations of military-related test situations requiring a degree of risk taking. The hypothesis tested was that women would be more conservative than men in performance involving risk.

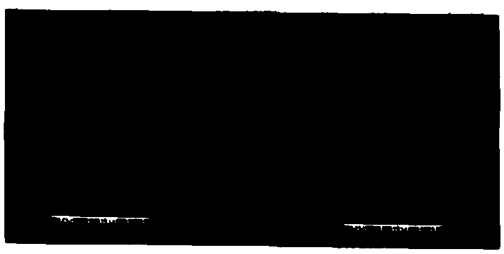
Methods

In the task shown in Figure 1, the subjects faced a screen displaying a simulated mine field with varying numbers and patterns of artillery-launched mines represented by dots in the field. The subjects were to decide whether or not to send a tank across the field based on their judgments of the chances of the tanks getting across successfully. Since the tank was not visible to the subjects prior to their decision, and since it could start from any point along the bottom line and proceed in a straight line through the field, the subjects had only the number of mines and their patterns as bases for a decision. They could decide "go" or "no-go" and were given points or lost points based on the outcomes of their decisions. The decisions made, the scores obtained and the times-to-decision were recorded automatically for each trial.

Figure 1

Examples of two of the patterns used are shown in Figure 2. The pattern on the left is one with a .90 probability of successful crossing. The pattern on the right has a probability of .30 for successful crossing. Twenty such randomly generated patterns were displayed for each probability level used.

Figure 2



The designs for the experiments are shown in Table 1. The subject groups were males and females. Subjects were tested only once in Experiment I, but over 4 test sessions in Experiment II. In Experiment I, the subjects were given extensive practice estimating the actual probability levels for successful crossing of the mine field prior to starting the decision-making phase described above. No such practice was given in Experiment II. In Experiment II, the easiest 20 trials, those with a probability of .90 for success, were eliminated to make the task shorter but more difficult.

Table 1

RISK-TAKING EXPERIMENTS

Experiment	Subject Groups	Test Sessions	Probabilities of Success	Trials
I	1. Males (N=18) 2. Females (N=18)	1. Probability Estimating 2. Decision Making (1 & 2 Same Day)	.90 .70 .50 .30 .10	20/Level of Probability
II	1. Males (N=9) 2. Females (N=9)	1. Decision Making 2. Decision Making 3. Decision Making 4. Decision Making (2/Day; Weeks 1 & 2)	.70 .50 .30 .10	20/Level of Probability

Results

The men and women did not differ significantly in their abilities to estimate probabilities of success in Phase I of Experiment I, or on their total scores for decision-making in either Experiments I or II.

In Experiment I, however, an analysis of the subjects' tendencies to choose the more conservative "no-go" response revealed that, for each set of 20 presentations of each field density, the males selected "no-go" a mean of 8.07 times to the females' mean of 7.03 times ($p<.05$). This difference held only for the more difficult middle range of densities (.70-.30) and not at the easier extreme densities (.90, .10). This finding appears to be contrary to the hypothesis that women would perform more conservatively than men.

Additionally, the women in this experiment tended to take longer than the men to make their decisions, particularly for the more difficult densities. This difference was significant for those fields with mine densities allowing for a .30 probability of successful crossing (decision time: male $\bar{X} = 2.56$ sec, female $\bar{X} = 3.76$ sec; $p<.025$).

When the experiment was partially replicated and extended to four sessions in Experiment II, the findings and interpretations became more complex.

Analysis of the subjects' tendencies to choose the "no-go" response again indicated that the males had a slightly (non-significantly) greater tendency than the females to choose the more conservative "no-go" response during the initial session (male $\bar{X} = 8.89$; female $\bar{X} = 8.61$). This pattern reversed for sessions 2-4; but, the sex X sessions interaction was not significant ($p<.25$). This reversal was great enough for the more dense fields, however, to result in a significant overall sex X field-density interaction ($p<.01$) as shown in Table 2. The prominent tendency for the women to choose the more conservative "no-go" response for the denser, more difficult fields supports the hypothesis that women would act more conservatively in situations involving risk.

Table 2

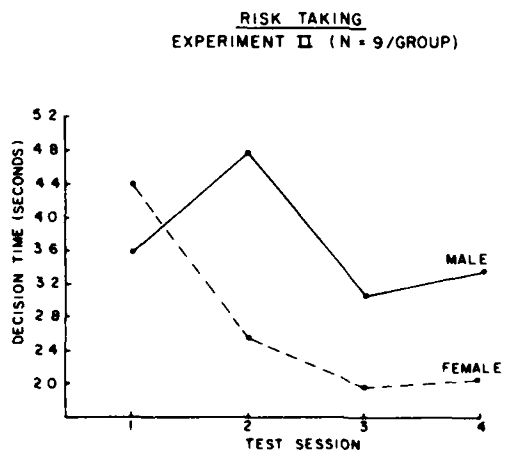
Risk Taking Experiment II

Mean Number "No-Go" Decisions (Out of 20)

GROUP	PROBABILITY OF SUCCESS .70	.50	.30	.10
Male (N=9)	1.31	4.42	13.14	18.47
Female (N=9)	0.72	5.36	16.17	19.64

Similarly complex patterns of results occurred for the decision-making latency measure in Experiment II. Figure 3 shows that, for the .30 probability density fields, a significant interaction occurred over test days. As in Experiment I, the women took longer than men to make their decisions on the first day of testing. However, on days 2-4 of testing, the women took significantly less time to make their decisions. The women appeared to have reduced their response time, while the men did not, over the 4 test days. *Although this interaction between sex and days was significant* only for fields with .30 probability densities, similar patterns occurred for the other probability densities as well in Experiment II.

Figure 3



Summary and Conclusions

1. While women were slightly less conservative in their performance on the initial session for the task described, they displayed a significantly greater tendency toward more conservative performance during the subsequent sessions. This predominantly conservative behavior for women in a risk-taking situation is consistent with the literature on decision making in risky driving situations where women drivers have been found to behave more conservatively than men drivers. Taken together, these results suggest that women may, in general, behave more conservatively than men in risk-taking situations.

2. The women took longer than men to make their decisions, particularly for the more difficult situations, during the initial session. However, this behavior pattern reversed for sessions 2-4 as had the tendency toward conservatism. This suggests that the more conservative response, once established, may involve less mental processing time.

3. The findings suggest that this type of computer simulation provides a reasonably valid and reliable means for studying risk-taking behavior in men and women.

4. These data clearly illustrate that one should be very cautious in drawing conclusions regarding male/female performance differences based solely on initial test trials. A very small amount of experience or training can have a dramatic effect on the relative performance of the groups. The value of obtaining repeated measures in sex-differences research is made quite apparent, as well.

References

Ebbesen, E. G., & Haney, M. Flirting with death: Variables affecting risk taking at intersections. Journal of Applied Social Psychology, 1973, 3, 303-324.

Hudgens, G. A., & Fatkin, L. T. Human performance: Psychological and physiological sex differences (A selected bibliography). Aberdeen Proving Ground, MD: US Army Human Engineering Laboratory, February 1980.

PANEL SESSION

THE EFFECTS OF PRESCRIBED MEDICATIONS ON COGNITIVE FUNCTIONING (CASE STUDY: IMPAIRED BRAIN FUNCTIONS DUE TO DIAZEPAM AND MEPROBAMATE

Al Brooker

Director, Neuropsych Services
David Grant Hospital, Travis AFB

The abuse of licit (prescribed) drugs continues to be a widespread problem, prevalent among adolescents and adults alike. According to Goodman and Gilman (1975, p 187), the two most widely used classes of prescribed drugs include the benzodiazepines (diazepam) and the propanediol carbamates (meprobamate). It has also been demonstrated that diazepam and meprobamate have variable effects; for example, some age-groupings are more sensitive to drug effects than other age-groupings (Howard, Hogan and Wright, 1975). Addiction to meprobamate can occur readily at as little as twice the usual daily dosage; high doses and prolonged exposure are assumed to be necessary to produce psysiologic addition to diazepam.

Diazepam is known to produce both physiological and psychological dependence and addiction. The latter is associated with withdrawal symptoms on abrupt cessation of the drug in certain predisposed individuals (Agrawal, 1978). Like meprobamate, diazepam depresses the duration of electrical aftercharge in the limbic system, including the septal region, the amygdala, and the hippocampus (Goodman and Gilman, 1975).

Meprobamate, like diazepam, is described quite commonly to allay anxiety. Our review of the literature did not allow us to draw any empirical conclusions regarding the drug's effect on human performance; i.e., there is inadequate dose-response data. Goodman and Gilman (1975) note that if a sufficiently large dose of meprobamate is given, psychological impairment will be found. Apparently, with the usual clinical dose of 400 mg. the effect of the drug is not reflected in performance on psychological tests. At a dose level of 1600 mg., however, there is definite impairment of learning, motor coordination and reaction time (Goodman and Gilman, 1975; Kometsky, 1958). Toxic reactions and side effects of meprobamate are sleepiness and ataxia and it is evident that tolerance and physical dependence will occur with meprobamate.

The case study presented demonstrated the important influence of diazepam and meprobamate on a patient's neuropsychological test performance. This presentation included: a brief medical history and summary of previous physical examinations; clinical interview and observations; brief description of the neuropsychological test battery; formulation of the neuropsychological findings; follow-up neuropsychological testing; and discussion. The patient was referred for assessment of cognitive, memory and motor functioning.

Brief Medical History: The patient (a 53 year old male) was admitted into a general hospital, approximately one month after the neuropsychological examination, for a gradually worsening syndrome over the past year, which had consisted of increasing confusion, episodes of slurred speech, and moderate difficulty with fine tasks and walking. Over the past year he had been seen in consultation by several neurologists on the assumption that he had some form of acute encephalopathic illness. He had also been seen in consultation by a psychiatrist who referred him for the neuropsychological examination. Detailed neurologic workup and heavy metal screen and CAT scan had failed to reveal any significant abnormality.

PANEL SESSION

PANEL SESSION:	THE WELL-BEING OF ARMY SOLDIERS AND THEIR FAMILIES: METHODOLOGICAL, SUBSTANTIVE, AND TECHNICAL CONSIDERATIONS
SESSION CHAIR:	Caren M. Carney (Walter Reed Army Institute of Research)
PANELISTS:	Terrence D. Fullerton, James A. Martin, and Richard Oldakowski (Walter Reed Army Institute of Research)
DISCUSSANTS:	Charlene L. Lewis and Edwin VanVranken (Walter Reed Army Institute of Research)

PROCEEDINGS ENTRIES

"The well-being of Army soldiers and their families: Methodological, substantive, and technical considerations" (Summary Paper)

"Comparing well-being and stress of several high risk Army groups" (Terrence D. Fullerton)

"Well-being and distress among drill sergeants: Civilian and military comparisons and the role of social support" (Caren M. Carney)

AD-P003 263

THE WELL-BEING OF ARMY SOLDIERS AND THEIR FAMILIES: METHODOLOGICAL, SUBSTANTIVE, AND TECHNICAL CONSIDERATIONS*

Caren M. Carney, Terrence D. Fullerton, Charlene S. Lewis, James A. Martin, Richard J. Oldakowski, and Ed Van Vranken

Department of Military Psychiatry
Walter Reed Army Institute of Research

Abstract

The Army's concern with health is shifting from a reactive and selective stance toward a proactive, preventive, and broadened orientation. Seven studies recently conducted in this department have used qualitative and quantitative methods to explore dimensions of well being and distress among soldiers and their families. Major concern has focused on the impact of group cohesion, social support, and socio-environmental conditions. Quantitative findings presented here illuminate comparative levels of well being and distress among over 3400 soldiers and spouses. The instrument used is the General Well-Being Schedule (Dupuy, 1978). Compared to civilian samples, only four of seven studies found overall General Well-Being scores in the range indicative of positive well being. All other samples fell into the moderately distressed range, although data collection is still in progress on the study of the spouses of the members of the Army's New Manning System COHORT units. Panel discussion focuses on correspondance between qualitative and quantitative observations.

Introduction

The social and behavioral sciences enjoy a sophisticated and appreciative audience among Department of the Army policy makers. Evidence of this sophistication can be found in the Army's willingness to view the soldier as an individual who occupies several roles and who moves within several social groups. Thus, the soldier may be a supervisor, a leader, a subordinate, a coworker, a "buddy," a spouse, a friend, and/or a parent. The Department of the Army has also developed an appreciation for the various stresses that can accrue from occupying such diverse roles and for the fact that competing roles may have negative consequences for the individual, the family, and group cohesion (Department of Defense, 1983).

Moving beyond a narrow view of stress, the Army is also encouraging the positive development and well being of soldiers and their families. This focus requires an understanding of the reciprocal effects of psychological and physical well being and distress. The interdisplinary perspective implied by these views presents both challenges and opportunities for those wishing to help map this largely unknown terrain.

The impetus for this panel derived from staff interest in the validity and generalizability of findings across studies recently conducted within our

*The views of the authors do not purport to reflect positions of the Department of the Army or of the Department of Defense.

Department. Investigators were particularly interested in assessing levels of psychological well being and distress as well as in determining the possible interrelations with physical and physiological states.

This presentation focuses on one instrument, The General Well-Being (GWB) Schedule, developed by Harold J. Dupuy (1978) and used in the National Center for Health Statistics'(1979) Health and Nutrition Examination Survey (HANES). Data from the HANES study allow comparisons with a national civilian sample; the comparisons are not the crucial issue here, they simply provide a convenient outside anchoring point. Discussions of the use and acceptability of the General Well-Being Schedule in the HANES study, as well as in other field studies, can be found in Fazio (1977), Husaini, Neff, and Stone (1979), and Neff and Husaini (1980).

The General Well-Being Schedule is designed to measure a construct which Dupuy (1978) names "general psychological well-being." The theoretical orientation driving this construct is that of a Lewinian field perspective. Thus, the construct is conceptualized as the "net impact of the many forces which affect an individual's subjective emotional or feeling states." (Dupuy, 1977, p. 1). The conceptualization of the construct is based on the assumption that individuals are capable of differentiating their feelings and of ordering these feelings along certain dimensions such as magnitude, intensity, and duration. The focus is on the person's own interpretation of internal, not external, conditions and states. Of no small difficulty (conceptual and semantic) is that Dupuy uses the term "well-being" to "reflect the affect quality of these states which may range from feeling distressed, disturbed, to feeling pleased, happy, or elated." (1977, p. 1). More research with Army service and family members should begin to provide some bases for comparisons and contrasts (especially normative ones) with civilian data.

Method

Eighteen items comprise the General Well-Being Schedule and they reflect six content areas addressing several health and mental health/disturbance issues (as distinguished from life satisfaction/happiness issues). Table 1 lists the content areas and the individual items comprising the schedule.

The distribution of the total GWB scores in the HANES study (N=6913) revealed a mean of 80.4 and a standard deviation of 17.4. The possible scores ranged from 0-110. The range of scores from 73-110 was indicative of positive well being, with the following internal breakdowns: elated (105-110), elevated (96-104), moderate-high (83-95), moderate-low (78-82), marginal (73-77). A range of scores from 61-72 indicated moderate distress, with similar breakdowns: discordant (69-72), upset (65-68), disturbed (61-64). Finally, severe distress was indicated by a range of scores from 0-60, broken down as follows: distraught (55-60), debilitated (49-54), impaired (37-48), disabled (18-36), depleted (0-17). Dupuy (1978) also provided estimates of the percentages of the civilian population falling into each category. For positive well being, the figure is 71.0%; for moderate distress, 15.5%; and for severe distress, 13.5%.

Data presented here derive from seven studies conducted in this department over the past several years. Interest has centered on several

TABLE 1

ADJUSTMENT FACTORS AND INDIVIDUAL ITEMS COMPRISING THE GENERAL WELL-BEING SCHEDULE OF DUPUY

A) <u>Freedom from Health Concern, Worry, Distress</u>:

Item # 10 Have you been bothered by any illness, bodily disorders, pains, or fears about your health?

Item # 15 How concerned or worried about your HEALTH have you been?

B) <u>Energy Level</u>:

Item # 9 Have you been waking up fresh and rested?

Item # 14 Have you felt tired, worn out, used-up, or exhausted?

Item # 17 How much ENERGY, PEP, VITALITY, have you felt?

C) <u>Satisfying, Interesting Life</u>:

Item # 6 How happy, satisfied, or pleased have you been with your personal life?

Item # 11 Has your daily life been full of things that were interesting to you?

E) <u>Cheerful Versus Depressed Mood</u>:

Item # 1 How have you been feeling in general?

Item # 4 Have you felt so sad, discouraged, hopeless, or had so many problems you wondered if anything was worthwhile?

Item # 12 Have you felt downhearted and blue?

Item # 18 How depressed or cheerful have you been?

F) <u>Relaxed Versus Tense, Anxious</u>:

Item # 2 Have you been bothered by nervousness or your "nerves"?

Item # 5 Have you been under or felt you were under any strain, stress, or pressure?

Item # 8 Have you been anxious, worried, or upset?

Item # 16 How relaxed or tense have you been?

G) <u>Emotional-Behavioral Control</u>:

Item # 3 Have you been in firm control of your behavior, thoughts, emotions, or feeling?

Item # 7 Have you had any reason to wonder if you were losing your mind, or losing control over the way you act, talk, thing, feel, or of your memory?

Item # 13 Have you been feeling emotionally stable and sure of yourself?

<u>Note</u>. Each item also asks "(DURING THE PAST MONTH)." Questions 1-14 have six structured response options ranging from none to an extreme amount of the attribute. Questions 15-18 have rating bars ranging from 0-10 with bipolar adjectives at each end only. The total score can range from 0-110 (see Fazio, 1977).

groups within the Army system perceived to be at greater risk for direct or indirect consequences of stress. Groups include: drill sergeants and non-drill NCO's, both members and spouses associated with the 82nd Airborne Division, members of the Special Forces (data for wives are not yet available), both members and spouses of the Department of the Army staff, and wives of the New Manning System COHORT units (data collection is still underway for this longitudinal study). Sample sizes are given in Table 2 and include breakdowns for major subgroups within the studies. Companion papers in this Proceedings by Carney and by Fullerton provide more detail on sample size and sample selection.

Results and Discussion

Table 2 provides mean GWB scores for the seven major study groups as well as for selected subgroups. A paper in this Proceedings by Fullerton analyzes group differences among these subpopulations.

Research investigation has also focused on qualitative appraisals of well being and distress among Army service personnel. Both qualitatively and quantitatively, discernible differences exist. Each group seems to possess a particular constellation of work, family, and individual assets and stressors that could mediate well-being or distress. For example, Special Forces

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR THE TOTAL GENERAL WELL BEING SCORE FOR DIFFERENT SAMPLES OF ARMY SERVICE AND FAMILY MEMBERS

SURVEY OR STUDY	MEAN	SD	N
Drill sergeants & NCO controls	68.5	19.1	2337
a) Drill sergeants	67.6	18.8	1190
b) NCO controls	69.4	19.4	1147
82nd Airborne members	69.5	18.9	423
Department of the Army staff	74.0	15.7	246
Special Forces members	76.7	16.9	132
82nd Airborne wives	73.7	18.7	81
Department of the Army wives	76.6	15.0	98
COHORT wives	63.6	19.1	121
a) 1st term wives	63.2	18.9	82
b) Cadre NCO wives	61.0	19.6	32
c) Officer wives	80.7	9.3	7

members possess high cohesion but exert considerable stress on family members because of the secrecy and danger in their jobs. Members of the 82nd Airborne Division also have dangerous missions and are often deployed from their families for long periods of time. Drill sergeants too are absent from their families. These NCO's report feeling isolated and under great job stress because of intense performance evaluations and long hours. Drill sergeants are similar to the Department of the Army (DA) staff in experiencing a low element of danger and long work hours. Unlike drill sergeants but similar to the Special Forces member, however, DA staff have high levels of perceived psychological control over themselves and their work. The General Well-Being Schedule is currently being administered to a scientific sample of wives married to enlisted service members assigned to New Manning System COHORT units.

The healthiest groups are those associated with greater perceived control and support (Special Forces, DA staff, and spouses of DA staff and 82nd Airborne members). Of concern are the negative long-term physiological consequences for individuals who experience chronic, moderate levels of stress. Our findings suggest that social support, high cohesion, and high levels of perceived control affect the net impact of negative stressors at the level of the individual, the family group, and the work group. Additionally, as a tool for psychiatric and behavioral epidemiology, the General Well-Being Schedule appears to be a reliable and valid measure of respondents' current emotional health. More extended discussions of the relevance of these findings for soldiers and their families, Army policy makers, clinicians, and the research community can be found in this Proceedings in papers by Carney and by Fullerton.

References

Department of Defense. Health promotion conference report, 20-24 June, 1983. [Report of a conference held in Reston, VA.]

Dupuy, H. J. (1978, October). Self-representations of general psychological well-being of American adults. Revision of paper presented at the meeting of the American Public Health Association, Los Angeles, CA.

Fazio, A. F. (1977, September). A concurrent validation study of NCHS' General Well-Being Schedule. Washington, D.C.: U. S. Government Printing Office (Superintendent of Documents, No. HE 20.6208.2/73).

Husaini, B. A., Neff, J. A., & Stone, R. H. (1979). Psychiatric impairment in rural communities. Journal of Community Psychology, 7, 137-146.

National Center for Health Statistics. (1979, March). Health and nutrition examination survey (HANES I) 1971-1975. Hyattsville, MD: National Center for Health Statistics. (NTIS No. PB-296 021).

Neff, J. A., & Husaini, B. A. (1980) Race, socioeconomic status and psychiatric impairment: A research note. Journal of Community Psychology, 8, 16-19.

Comparing Well-Being and Stress of Several High Risk Army Groups

Terrence D. Fullerton
Walter Reed Army Institute of Research

Abstract

The investigators of the Department of Military Psychiatry explored the well-being and distress of soldiers in a variety of military groups, some of which are considered high risk populations. Soldiers of several of these groups seem to possess somewhat different work, family, and individual assets and stressors. These groups included drill sergeants, soldiers deployed to the Sinai as part of the multinational peacekeeping force, Special Forces soldiers, and members of a Department of Army (Pentagon) staff. The general psychological well-being of these soldiers was assessed by the General Well-Being Schedule (Dupuy, 1978). The scores on the General Well-Being Schedule of all but two of the groups are in the moderately distressed range. The scores of the Special Forces soldier and members of the DA staff are in the marginally positive range of well-being. There are significant group differences on the overall and sub-scales of the General Well-Being Schedule.

Introduction

Stress and well-being have become topics of major interest to Department of Army policy makers and many major Army commanders. This interest seems to focus on determining: what are the major stresses experienced by their soldiers, which are causing the greatest dysfunction, and how to ameliorate the stressful effects. Investigators of the Department of Military Psychiatry have recently explored the well-being and distress of several military groups, some of which seem to be in stressful environments. Soldiers of the 82nd Airborne Division conduct somewhat dangerous missions (e.g., Grenada) and repeatedly deploy from their families for several weeks each month. In 1982 one battalion of the 82nd Airborne Division was deployed to the Sinai and was positioned between the Israeli and Egyptian armies. Soldiers in this battalion performed as members of the first Multinational Force and Observers in the Sinai for six months. The deployment of soldiers is typically accompanied by exposure to environmental and psychosocial stressors. Novel stressors in this deployment included uncertainty about: the possibility of conflict between these opposing armies, the living conditions, their reactions to this novel role as peacekeepers, the heat, and later, boredom.

The Special Forces (SF) soldiers are in some of the most frequently deployed units. Although a six-month deployment for the 82nd troops is a rarity, these SF soldiers frequently deploy for extended periods, often on real-world missions to train and advise soldiers in environments often perceived as dangerous (e.g., El Salvador, Honduras, Lebanon, and several African nations). On the other hand, these soldiers belong to what is widely considered to be one of our most cohesive groups. These soldiers are reputedly closely bonded to other members of their units and spend considerable on- and off-duty time "team building."

Drill sergeants perceive themselves and are seen by others as one of the most stressed groups in the Army. Army folklore suggests that basic training is often the most stressful environment a soldier may be subjected to in the course of his career, short of war. Initially, researchers focused on the stresses of trainees (e.g., Datel, Gieseking, Engle, and Dougher, 1966; Datel & Engel, 1966; Datel & Lifrak, 1969) but recently have also focused on the considerable stresses experienced by the drill sergeants. Drill sergeants report feeling isolated and working extremely long hours under constant, close, and critical supervision. The Raupp report, TRADOC's study of basic training, suggests that the stress and strain incurred by drill sergeants may increase the possibility of trainee abuse while decreasing the level of trainee effectiveness. A companion paper in this Proceedings (Carney, 1984) focuses on the well-being and distress among drill sergeants and non-drill NCOs, particularly with reference to the impact of social support.

Members of the Department of Army (Pentagon) staff also view their organization as representing a particularly stressful environment. The officers are older than the soldiers of the other groups. These staff officers report working long hours (nine or more); feeling "run-down" by the end of the week; experiencing continuous job stress due to unrealistic deadlines, competing tasks, and multiple bosses; and being dissatisfied with the amount of time and energy left for family and self (Marlowe, 1983). On the other hand, they report satisfaction with the way their expert knowledge is used, with the trust and respect shown to them by their superiors, and with the support given by superiors and co-workers.

Most of these groups report high perceived stress, and somewhat different work, work group, and family assets and stressors. Thus, this research examines the well-being and distress among various separate groups of soldiers while considering the impact of group cohesion/social support and socio-environmental conditions. This information is important for policy makers for knowing whether/how to ameliorate stress and moving toward the preventative focus the Army is now pushing.

Method

Subjects

The subjects were 3228 members of the United States Army. Sub-groups included: 437 soldiers of a battalion of the 82nd Airborne Division; 139 Special Forces soldiers; 1216 drill sergeants; 1179 noncommissioned officers stationed at the same training posts as the drill sergeants; and 253 members of a Department of Army staff.

Data Collection Procedures and Instrument

The General Well-Being Schedule was one of an extensive battery of instruments examing the amount of stress experienced on the job and at home by these soldiers. The General Well-Being Schedule is an eighteen item instrument designed to measure the net result of the various factors which affect an individual's general psychological well-being (Dupuy, 1978). The possible scores range from 0-110. Dupuy suggests the following break down of scores: (73-110) indicates positive well-being; (61-72), moderate distress; and (0-60), severe distress.

In addition to an overall GWB score, there are six subscales addressing several health/mental health issues. A more detailed description of the GWB and the subscales is available in this Proceedings in a paper by Carney, Fullerton, Lewis, Martin, Oldakowski, & Van Vranken (1984), Fazio (1977) and Neff & Husaini (1980).

Results and Discussion

Dupuy (1978) suggested that there were six subscales in the General Well-Being Schedule. Factor analysis of the entire military data base replicated to a large extent (five of the six factors matched perfectly) Dupuy's subscales. Cronbach's (1951) coefficient alpha, an index of internal consistency, for the subscales varies from .46 to .82.

The next set of analyses allowed us to examine group differences of the overall General Well-Being Scale and each of the six subscales scores. There are significant differences between the groups on the scores of the overall GWB and each subscale. The results of Duncan's multiple range tests are presented in Table 1. These results indicate that the drill sergeants, NCOs at basic training posts, and the soldiers of the peacekeeping force all have GWB scores in the moderately distressed range and the scores of these three groups are lower than the scores of the Special Forces soldiers and the DA level staff. A consistent pattern of results emerged. SF soldiers reported the most positive scores, followed by members of the DA staff, and with the NCOs, Drills, and 82nd soldiers slightly shifting order among themselves.

Explanations of these data may lie with the somewhat different work, work group, family, and environments characteristic of these groups. The SF has high cohesion (work group social support) but dangerous missions and family stress due to the danger, frequent extended deployments and secrecy. The soldiers of the 82nd also have dangerous missions, frequent deployments but less cohesion/work group support (Fullerton & Manning, 1984). The drill sergeants have little physical danger, low work group support, perceived lack of control, extended hours, and evaluation stress. The NCOs are similar to the drill sergeants but with less evaluation stress. The members of the Army staff are similar to the drill sergeants in encountering little physical danger, extended work hours, and high evaluative stress. Unlike drill sergeants but similar to the SF, the DA staff has work group support and perceived control over themselves and their work environment. Thus the healthiest groups are those with the greatest work group support and perceived control. The possible long-term negative consequences of the moderate distress on the physical and mental health of these drill sergeants, NCOs, and the soldiers of the 82nd are disconcerting. Follow-up research tracking the long-term consequences, the positive impacts of work group support, and the efficacy of interventions is needed.

The views of the author do not purport to reflect positions of the Department of the Army or the Department of Defense.

REFERENCES

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Datel, W. & Engle, E. Affect levels in another platoon of basic trainees. Psychological Reports, 1966, 19, 903-909.

Datel, W., Gieseking, C., Engle, E. & Dougher, M. Affect levels in a platoon of basic trainees. Psychological Reports, 1966, 18, 271-285.

Datel, W. & Lifrak, S. Expectations, affect change, and military performance in the Army recruit. Psychological Reports, 1969, 24, 855-879.

Dupuy, H.J. (1978, October). Self-representations of general psychological well-being of American adults. Revision of paper presented at the meeting of the American Public Health Association, Los Angeles, CA.

Fazio, A.F. (1977, September). A concurrent validation study of NCHS' General Well-Being Schedule. Washington, D.C.: U.S. Government Printing Office (Superintendent of Documents, No. HE 20.6208.2/73).

Fullerton, T.D. & Manning, F.J. Effects of eliteness - a matter of cohesion or self-selection. Paper submitted for presentation to the 1984 Annual Convention of the APA, Toronto, Ontario, Canada, August 1984.

Marlowe, D.H., Camp, N.M. & Jones, F.D. Interim Report ODCSPER Stress Management/Corporate Fitness Study, 1983.

Neff, J.A. & Husaini, B.A. (1980) Race, socioeconomic status, and psychiatric impairment: A research note. Journal of Community Psychology, 8, 16-19.

Raupp. Task Force on Initial Entry Training Leadership (1978, December). Toward positive leadership for initial entry training. Fort Monroe, VA: U.S. Army Training and Doctrine Command.

TABLE I
Duncan's Multiple Range Test
General Well-Being Scale
(* denotes pairs of groups significantly different at .05 level)

Overall General Well-Being (Range 0-110)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							11.39	.0000
67.59	Drill							
69.43	NCO	*						
69.53	82							
73.97	DA Staff	*	*	*				
76.66	SF	*	*	*				

Freedom from Health Concern, Worry Distress (Range 0-10)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							10.23	.0000
10.22	Drill		*	*				
9.40	NCO							
9.56	82							
10.41	DA Staff		*	*				
10.48	SF		*	*				

Energy Level (Range 0-15)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							28.42	.0000
10.02	Drill							
11.33	NCO	*						
11.32	82	*						
11.83	DA Staff	*						
13.12	SF	*	*	*	*			

Satisfied, Interesting Life (Range 0-10)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							15.16	.0000
5.00	Drill							
5.45	NCO	*		*				
4.95	82							
5.88	DA Staff	*	*	*				
6.11	SF	*	*	*				

Cheerful versus Depressed Mood (Range 0-20)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							12.36	.0000
15.94	Drill							
16.03	NCO							
15.84	82							
17.75	DA Staff	*	*	*				
18.04	SF	*	*	*				

Relaxed versus Tense, Anxious (Range 0-20)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							9.43	.0000
14.42	Drill							
15.25	NCO	*						
15.82	82	*						
15.44	DA Staff	*						
16.40	SF	*	*					

Emotional Behavioral Control (Range 0-15)

Mean	Group	Drill	NCO	82	DA	SF	Overall F	Overall P
							2.91	.02
12.06	Drill							
11.93	NCO							
12.06	82							
12.48	DA Staff	*	*					
12.51	SF		*					

AD-P003 265

WELL BEING AND DISTRESS AMONG DRILL SERGEANTS: CIVILIAN AND MILITARY COMPARISONS AND THE ROLE OF SOCIAL SUPPORT*

Caren M. Carney

Department of Military Psychiatry
Walter Reed Army Institute of Research

Abstract

This study compares civilian and military scores on the General Well-Being (GWB) schedule of Dupuy (1978). Civilian data derive from a national sample of American adults interviewed during the National Center for Health Statistics' (NCHS) Health and Nutrition Examination Survey (HANES), wave 1 (NCHS, 1979). Military data derive from a study of drill sergeant stress which surveyed drill and non-drill Non-commissioned officers (NCO's) from the same training posts (Vernon, Marlowe, Datel, & Holloway, 1980).[1] Military scores on the GWB were lower than civilians' scores; drill sergeants had lower scores than non-drill NCO's. In both studies, high levels of social support had pervasive direct ameliorating effects on distress.

Introduction

The U. S. Army has increasingly expressed interest in exploring new avenues for achieving an optimal basic training environment. The "Raupp Report" attempted a comprehensive view of the leadership climate in Army training centers and the way this climate affected both new recruits and drill sergeants (Task Force on Initial Entry Training Leadership, 1978). The new recruit is commonly viewed as the individual under intense stress, and indeed the report listed 16 stressors new recruits experience often. However, the report also listed 27 stressors that affect drill sergeants. These ranged from long hours, being drafted to become a drill sergeant, and concern about the career consequences of withdrawing from the program, to shortages of drill sergeants, unrealistic training schedules, and lack of privacy. The potential for such stress to result in trainee abuse was a logical and real concern.

The Army's concern with the stressors affecting drill sergeants reflects an awareness that successful teaching requires both a healthy learning environment and a healthy teacher. The concern also reflects an appreciation that the drill sergeant not only teaches specific skills, but also imparts specific values and standards. The drill sergeant also often provides new

---

*The views of the author do not purport to reflect positions of the Department of the Army or of the Department of Defense.

[1]Darleen M. Vernon, CPT, MSC, now at the University of Pittsburgh, developed and executed this study and served as its project officer. Richard J. Oldakowski played a crucial role in coordinating and conducting the complex data-base management aspects of this study. Thanks also extend to David H. Marlowe for providing essential guidance and to other Departmental staff who assisted on this study. Finally, grateful appreciation is extended to our study participants whose cooperation made this research possible.

recruits with their first role model of the "ideal" soldier. Thus, the levels of well-being and distress among drill sergeants are bound to be reflected, respectively, in better or poorer effectiveness in leading, guiding, and teaching the new recruit how best to be a competent soldier.

Novaco, Sarason, Robinson, and Parry (1983) suggest that for Marines, the drill sergeant school selection process tends to graduate candidates who are least at risk for stress (compared to those who drop out or are not graduated for psychological/behavioral reasons). It is uncertain whether the screening process operates similarly in the Army. If it does, then stress levels among drills would require scrutiny in the context of their better coping skills at the onset of their duties. This study attempts a first step in determining whether drill sergeants are more stressed than their non-drill counterparts and suggests the importance of social support in mitigating distress.

Method

Participants in the study were 2405 of 5000 NCO's sampled at eight Army training posts. Most respondents were male, with 1223 (49.3%) identifying themselves as drill sergeants. Both drill and non-drill groups answered identical questionnaires which asked for demographic information and social support ties as well as for information on physical and psychological well being and on satisfaction with the Army (see Vernon et al., 1980).

The HANES study used a sample of 6913 Americans to assess physical and psychological status (NCHS, 1979). The GWB Schedule was one of the instruments used to assess mental health (Dupuy, 1978; also, see a paper in this Proceedings by Carney, Fullerton, Lewis, Martin, Oldakowski, and Van Vranken). To provide another comparison group with the NCO study, a subsample of men was chosen from the total HANES population. "Working men" were defined as those men between 25 and 45 years of age, having a high school or college degree only, and possessing a history of paid employment either as a self-employed worker or as a private or government employee.

Social support was assessed in the same way for both groups by a social support question included as part of the GWB interview schedule in the HANES study. Identification as an individual with low versus high support was based on the response to the question, "Do you discuss your problems with any member of your family or friends?" High support answers included: "No, I do not have any problems," "Yes, and it helps a lot," or "Yes, and it helps some." Low support answers were: "Yes, but it does not help at all," "No, I do not care to talk about my problems with anyone," "No, no one cares to hear about my problems," or "No, I do not have anyone I can talk with about my problems." Thus, this measure of social support taps aspects of the individual's actual social network as well as perceptions of social support. Excluding from analyses those individuals who answered "No, I do not have any problems" did not substantially change the high support group means. Including this response category did substantially raise the low support group means. This response category could be considered an index of well-being but is included in the high support group due to the question's context.

Results

For both the HANES study and the NCO study, social support played a powerful role in mitigating distress. In the HANES study, the Pearson

correlation between the continuous social support measure and the total GWB score was .409 (all correlations discussed here are significant beyond the .001 level). The correlations for subscales 1 through 6 respectively were .241, .296, .328, .386, .379, and .305. For the working men, the total and subscale correlations were, respectively: .398, .147, .267, .316, .409, .359, and .243. Also, compared to the rest of the HANES study population, they had significantly higher GWB scores (see Table 1 for mean effects). (T tests were all significant beyond the .001 level for the total and subscale scores).

A key question in the NCO stress study concerned how social support and occupation might explain well-being/distress. Several possible explanations exist: neither contributes additively, one or both contribute additively, or the synergistic or interactive combination of the two either does or does not contribute to explaining well-being/distress. This last possibility represents the concept of interaction (or non-additivity) wherein the effect for one independent variable is not the same under all conditions of the other independent variable. Thus, e.g., given the presumedly high levels of stress for drill sergeants, would high levels of social support mitigate the effects of stress only for controls (i.e., and not for drill sergeants)? If so, an interaction measure would uncover that possibility.

Multiple regression was used to explore the possible contributions that social support and occupation make to well-being/distress. This procedure uses a multiple correlation coefficient (R) to assess the variation that two or more independent variables share with a single dependent measure. Here the three independent variables are occupation, social support, and a cross-product of the standardized measures of occupation and social support. Seven multiple regressions were computed, one each for the GWB total and 6 subscale (dependent) measures. The respective R's for the set of 3 independent variables were: .452, .190, .413, .436, .429, .386, and .348 (all were significant beyond the .001 level).

Given this 3 variable model, it is also useful to ask which of the independent measures accounted for substantial amounts of variance. Inspecting the partial correlation coefficients provides a means for assessing the contribution of any one independent variable while statistically holding the effects of the others constant.

Similar to an ANOVA approach, it is appropriate first to inspect for possible interaction effects. For no one of the dependent measures did the interaction account for even 1% of the variance. Next, inspecting the partial additive contribution of the occupation measure uncovered few major differences between drills and non-drills. Occupation contributed 3% to explaining energy level and 1% to explaining subscales 1, 3, and 5. (See Table 2 for specification of mean effects).

For each dependent variable, the largest contribution to the set of 3 independent measures consistently came for the additive measure of social support. The partial correlations of social support with the total and subscale scores were, respectively: .448, .162, .388, .428, .427, .376, and .345 (all highly statistically significant). Comparing these partials with the R's above, it is clear that most of the variance is explained by the direct effects of social support. The respective Pearson correlations are: .447, .163, .380, .425, .426, .373, .346.

TABLE 1

GENERAL WELL-BEING SCHEDULE MEANS AND STANDARD DEVIATIONS FOR THE TOTAL SAMPLE AND SELECTED SUBPOPULATIONS IN THE HEALTH AND NUTRITION EXAMINATION STUDY

GWB Component:	Total Score	Subscale 1	Subscale 2	Subscale 3	Subscale 4	Subscale 5	Subscale 6
Description:	Distress Versus Well Being	Freedom From Health Concern, Worry	Energy Level	Satisfying, Interesting Life	Cheerful Versus Depressed Mood	Relaxed Versus Tense, Anxious	Emotional-Behavioral Control
Range of Scores:	(0-110)	(0-15)	(0-20)	(0-10)	(0-25)	(0-25)	(0-15)
Sample/subsample:							
Total:	80.31 (17.67) [6913]	11.13 (3.82) [6913]	13.04 (4.13) [6913]	6.70 (2.06) [6913]	18.52 (4.52) [6913]	17.97 (5.19) [6913]	12.95 (2.30) [6913]
Low Support:	71.68 (20.17) [1071]	9.85 (4.37) [1071]	11.44 (4.54) [1071]	5.77 (2.31) [1071]	16.33 (5.23) [1071]	16.16 (5.94) [1071]	12.12 (2.88) [1071]
High Support:	81.90 (16.70) [5842]	11.37 (3.66) [5842]	13.33 (3.98) [5842]	6.87 (1.97) [5842]	18.92 (4.26) [5842]	18.30 (4.97) [5842]	13.10 (2.14) [5842]
Working men:	85.56 (13.22) [742]	12.14 (2.95) [742]	14.30 (3.32) [742]	7.04 (1.83) [742]	19.89 (3.39) [742]	18.58 (4.36) [742]	13.60 (1.67) [742]
Low Support	78.50 (16.02) [84]	11.27 (3.32) [84]	13.31 (3.82) [84]	6.08 (2.06) [84]	18.01 (4.16) [84]	16.86 (4.99) [84]	12.96 (2.25) [84]
High Support	86.46 (12.55) [658]	12.25 (2.88) [658]	14.43 (3.23) [658]	7.17 (1.77) [658]	20.13 (3.21) [658]	18.80 (4.22) [658]	13.68 (1.56) [658]

Note. The figures given represent the mean, standard deviation (in parentheses), and sample size (in brackets). Higher scores on the GWB indicate better adjustment. Analyses were conducted by the author using data from the National Center for Health Statistics (NCHS, 1979). See text for discussion of the derivation of social support dichotomy and for description of sample of working men.

TABLE 2

GENERAL WELL-BEING SCHEDULE MEANS AND STANDARD DEVIATIONS FOR THE TOTAL SAMPLE AND SELECTED SUBPOPULATIONS IN THE DRILL SERGEANT STRESS STUDY

GWB Component:	Total Score	Subscale 1	Subscale 2	Subscale 3	Subscale 4	Subscale 5	Subscale 6
Description:	Distress Versus Well Being	Freedom From Health Concern, Worry	Energy Level	Satisfying, Interesting Life	Cheerful Versus Depressed Mood	Relaxed Versus Tense, Anxious	Emotional-Behavioral Control
Range of Scores:	(0-110)	(0-15)	(0-20)	(0-10)	(0-25)	(0-25)	(0-15)
Sample/subsample:							
Total:	68.49 (19.14) [2337]	9.82 (3.90) [2392]	10.66 (4.41) [2385]	5.22 (2.49) [2396]	15.98 (5.08) [2360]	14.83 (5.44) [2380]	12.00 (2.86) [2390]
Non-drill NCO's:	69.43 (19.42) [1147]	9.40 (3.94) [1177]	11.33 (4.35) [1171]	5.44 (2.46) [1178]	16.03 (5.24) [1161]	15.25 (5.49) [1171]	11.93 (2.95) [1174]
Drill Sergeants:	67.58 (18.82) [1190]	10.02 (3.81) [1215]	10.22 (4.37) [1214]	5.00 (2.50) [1218]	15.93 (4.92) [1199]	14.42 (5.37) [1209]	12.06 (2.78) [1216]
Low Support:	56.73 (18.17) [644]	8.88 (4.02) [659]	8.37 (4.09) [659]	3.79 (2.29) [660]	13.00 (5.02) [651]	12.07 (5.15) [656]	10.57 (3.32) [659]
High Support:	72.97 (17.51) [1686]	10.17 (3.78) [1721]	11.53 (4.21) [1717]	5.76 (2.35) [1724]	17.12 (4.61) [1702]	15.87 (5.17) [1713]	12.54 (2.45) [1720]

Note. The figures given represent the mean, standard deviation (in parentheses), and sample size (in brackets). Higher scores on the GWB indicate better adjustment. See text for discussion of the derivation of the social support dichotomy and for description of non-drill sergeant and drill sergeant subsamples.

Discussion

This research found few real differences between drill and non-drill NCO's. This surprising non-finding counters the stereotype of the stressors of drill sergeant duties as second only to those of combat. Do drills cope more effectively by complaining? Are all NCO's at occupational risk for stress? The NCO role does appear analogous to that of the mid-level manager who experiences demands both from below and above. Of concern here are the substantially lower levels of well-being and distress among the NCO's compare to American adults and particularly to working men. Or is it the case that the training posts exert high stress not only on the drill sergeant, but also on the non-drill NCO counterpart? Further research will be required to point conclusively to this last possibility as the best explanation.

It is true that social support does improve well-being for NCO's in the same way that it worked for adults in the HANES study. Clearly it exerts powerful direct effects. However, the level differences between the Army and civilian groups are striking. These will require more analyses and discussion than space permits here. Interested readers can make comparisons now, however, given the information presented in the tables.

Campbell (1981) has suggested using a conceptual model of well-being that considers positive and negative well-being as distinct entities. He also argues that life satisfaction (or social well being), though conceptually distinct, should be related to individual well-being and distress. Future analyses will employ factor analytic techniques to determine the adequacy of this conceptualization and to assess correspondence with factors found by Dupuy (1978) in the HANES study.

References

Campbell, A. (1981). The sense of well-being in America: Recent patterns and trends. New York: McGraw-Hill.

Dupuy, H. J. (1978, October). Self-representations of general psychological well-being of American adults. Revision of paper presented at the meeting of the American Public Health Association, Los Angeles, CA.

National Center for Health Statistics. (1979, March). Health and Nutrition Examination Survey (HANES 1) 1971-1975. Hyattsville, MD: National Center for Health Statistics. (NTIS No. PB-296 021). [Document plus data tape].

Novaco, R. W., Sarason, I. G., Robinson, G. L., & Parry, C. (1983). Attributes of drill instructor school graduates: Stress-related factors. Arlington, VA: Office of Naval Research. (Technical Report DI-ONR-001).

Task Force on Initial Entry Training Leadership. (1978, December). Toward positive leadership for initial entry training. Fort Monroe, VA: U. S. Army Training and Doctrine Command.

Vernon, D. M., Marlowe, D. H., Datel, W. E., & Holloway, H. (1980, May). Drill sergeant and NCO stress and its consequences for performance in the training environment: A comparative study. [Research protocol]. Washington, D. C.: Walter Reed Army Institute of Research.

AD-P003 266

HABITABILITY AND HUMAN PRODUCTIVITY ISSUES CONFRONTING THE AIR FORCE SPACE COMMAND

Jack Stuster, Ph.D.
Senior Scientist
Anacapa Sciences, Inc.

Abstract

There are four areas of planned and proposed DoD space activity in which habitability issues play an important role. These are: (1) long-duration space missions, (2) short-duration space missions, (3) extended-duration STS missions, and (4) terrestrial space operations center(s). The purpose of this paper is to identify some of the missions within each of these areas and to explore the associated behavioral, psychological, and sociological issues. These include: sleep, clothing, exercise, medical support, personal hygiene, food preparation, group interaction, habitat aesthetics, outside communications, recreational opportunities, privacy and personal space, and waste disposal and management.

INTRODUCTION

There are four areas of planned and proposed DoD space activity in which habitability issues play an important role. These are: (1) long-duration space missions, (2) short-duration space missions, (3) extended-duration STS missions, and (4) terrestrial space operations center(s). The following paragraphs provide brief discussions of the principal habitability and human performance issues.

Before we proceed, a few definitions are in order. **Habitability** may be operationally defined as those aspects of an environment that either immediately or in the long run affect human performance and productivity (both positively and negatively). This definition acknowledges both degrading and enhancing elements. The objective of understanding habitability issues is to design environments and associated operations in order to facilitate performance and to optimize sustained human productivity. This objective is achieved by the application of all available expertise to the development of effective design and operations principles. Remember, productivity represents the ultimate objective of our efforts. **Short-duration missions** are missions no longer than 14 days. This includes all planned STS missions within current orbiter capabilities. **Long-duration missions** are those lasting longer than 60 days. Proposed space station tours are within the 60-90 day range. Recent suggestions to conduct **extended duration,** 30-day DoD STS missions complicate the matter. That is, 30 days aboard an orbiter would substantially elevate the importance of STS habitability issues--the vehicle simply was not designed, in terms of habitability, for routine 30-day tours with full crews. Performance on 30-day STS missions may be more seriously affected than on 90-day tours aboard a space station designed for long-duration missions. Consequently, extended-duration STS missions must be considered separately from both short duration shuttle missions and long-duration sojourns aboard a space station.

Issues of habitability are important during short-duration missions such as shuttle flights. They are also important under vigilant or emergency conditions that must be anticipated for survivable ground control centers in both routine and post-attack environments. Clearly, though, habitability issues are **most** salient in the area of long-duration space missions. For this reason, our discussion will focus on the behavioral requirements associated with long-duration missions.

OPERATIONS PERFORMANCE REQUIREMENTS

It is generally recognized that many individuals are both willing to and capable of performing adequately under adverse conditions when those conditions are of short duration, extreme novelty, and when the rewards for enduring the conditions are significant. However, when humans are subjected to adverse conditions for long durations, on a routine basis and in the absence of substantial reward or recognition, performance degradation **must** be expected.

We are not concerned here with system tests or even with the first several missions of a 90-day-tour space station. Those crews would likely perform admirably under extremely inhospitable living conditions. Conversely, what we must design from the start is a system that would be not only acceptable but also conducive to work under routine conditions--during the 35th mission or during an individual crew's **tenth** tour on a rotational basis. It is imperative that we design truly functional habitats--environments that both support life and encourage productivity--if future space missions are to be successful. It is equally important that habitability and human performance issues associated with ground control facilities are addressed in a systematic fashion.

There are several behavioral issues associated with proposed long-duration missions that will inevitably affect human performance; these issues are discussed below. In identifying these issues we take the initial step toward avoiding problems and mitigating effects. At the very least, once identified, problems can be anticipated. Essentially, we need to know what factors characteristic of confinement and isolation contribute to the degradation of human productivity. We need to explore these issues in order to design systems that avoid or minimize the negative effects on the human component. If isolation is, as Carlyle has described it, "The sum total of wretchedness to man," how may we render it less wretched?

ISSUE IDENTIFICATION

We have limited our concern, at this early stage of concept development, to only those biological, psychological, and sociological issues with architectural or hardware design implications. While other issues, such as crew selection and evaluation, are obviously important, it is critical to first explore those concerns which demand long lead times for research and development, and those that require early architectural definition. Identification of these requirements during the early stages of design is imperative in order to maximize the productivity on future long-duration space missions and to minimize system costs. These potential costs include both operational compromise and expensive retrofitting to mitigate design flaws.

The issues we have identified are addressed in subsequent paragraphs, including discussion of possible solutions and research requirements.

Sleep. The variables associated with sleep include schedule, total amount per cycle, and "quality" (quality is defined as time to onset and number of awakenings). Variation in each of these dimensions affects human performance. We need to discover means to assess these variables and to apply the lessons operationally. Questions of shift work rather than synchronous schedules are particularly important since they impinge on other issues such as the maintenance of privacy and personal space. Research is required in this area to define an appropriate compromise between operational and human requirements.

Clothing. Much has been learned from Skylab and STS missions regarding clothing [illegible] a "shirt sleeve" working environment in space. Air Force concerns about clothing [illegible] satisfied by closely monitoring the NASA experience. Also, specific aspects of [illegible] missions may require the development of specific items, primarily in [illegible] protective/shielding equipment and/or fabrics.

Exercise. Exercise is a critical concern on long-duration space missions. Due to the many effects of muscle atrophy and the particularly insidious bone decalcification characteristic of microgravity environments, special attention must be given to defining specific forms and amounts of exercise. Many of the required data are already available from previous space missions, in particular Skylab. Further work is required, however, in order to address the psychological aspects of exercise. That is, what may be required is the development of methods of exercise that are less boring and more motivating than those that have been employed or proposed for astronaut use. Attempts must be made to render exercise more recreational; for instance, place the ergonometer in front of a window.

Medical Support. The importance of an onboard medical support capability increases with mission duration. No serious medical emergencies have been reported during American space missions. The Soviets have not, apparently, been so fortunate. Evidence from Sealab, a highly relevant analogue to space missions, indicates that all similar ventures should involve close coordination of medical and engineering phases with a great deal of control vested in the medical complement of the team. The degree to which this suggestion is possible in an evolving system requires further exploration. Similarly, a paramount fear expressed by Valery Ryumin during his 175 day sojourn aboard Salyut-6 was that he might experience an attack of appendicitis or an abscessed tooth. Fear of appendicitis or deep dental pain could manifest itself in several forms, each of which could affect productivity. Preventive measures for both emergencies and the **fear** of medical emergencies could be quite simple (e.g., require that each long-duration astronaut have his or her appendix removed before flying, and provide dental emergency kits for onboard use). The concern for treating accidental injury, however, requires further attention. Additional work is also required to identify medical concerns from the perspective of crew personnel.

Personal Hygiene. Hygiene parameters for long-duration space missions are likely to differ substantially from what is acceptable during short forays into space. While "camping out" conditions pose less of a threat during short duration missions and during emergencies, "camping out" can be expected to seriously degrade human performance under routine conditions during long-duration space missions. During the course of the space program, improvements in onboard hygiene technology have been substantial. Further improvements are necessary for onboard systems to be acceptable for long-duration missions. What appears to be required is greater sensitivity to behavioral and psychological factors affecting user acceptability.

Food Preparation. Perhaps the quintessential habitability issues involve food. There is much to learn from analogous experiences such as submarines and off-shore oil platforms concerning the importance of quality food in maintaining high morale and productivity. It must be remembered that we are considering missions of both long-duration and a routine nature. Under such conditions attitudinal problems regarding "inadequate" food may be inevitable regardless of personal motivation, loyalty, and the like. Certain earthbound and on-orbit experiences suggest that the quality, variety, and availability of food represents one of the fundamental issues affecting human performance on long-duration missions. Nutritional content of meals is not in question. What requires further definition is the optimum compromise between quality and operational conditions. Further studies may indicate that this is one area where little compromise is possible with behavioral principles.

Group Interaction. Under this topic must be considered two separate issues: organizational structure **and** communication. In terms of organizational structure the form that has evolved within NASA for STS missions (quasi-military organization with mission specialists and payload specialists) is consistent with all other smoothly functioning analogous conditions. While it is expected that this will be the social structure of

a NASA space station, some observers have speculated that scientists and technicians may object to a quasi-military atmosphere and require an absence of onboard central authority. Others have even assumed that the organization would (or should) be egalitarian and that crew size should be an odd number to disallow stalemated votes. Clearly, it is silly to consider a town meeting approach to space station management. A more appropriate analogy might be a research ship and, later in the station's evolution, a distant-water fishing vessel (i.e., mission and production orientation). Research is required to explore the dynamics of crew interpersonal relations in order to identify appropriate levels of participation in onboard decision-making. Also required is the development of procedures for establishing rules and regulations aboard semi-autonomous spacecraft.

In terms of effective communications within the crew, it has been indicated that an onboard space large enough for all individuals to meet is required; it is typically the most heavily used space under analogous conditions of isolation and confinement (e.g., Sealab, Tektite). Such a space--whether called the bridge or the wardroom--may need to accommodate both outgoing and incoming crews and perhaps even the transport crew for short durations. Further attention to defining wardroom requirements is necessary; possible solutions involve use of flexible internal structures. In addition to group meetings, provisions must be made to allow both routine and emergency communications among the crew (e.g., notify of minor anomalies at watch change, bulletin-board information, general quarters). A complete analysis of onboard communications requirements is necessary in order to develop solutions that allow for safe and efficient exchanges of information among the crew.

Habitat Aesthetics. Much intuitive and some objective evidence supports the notion that the "visual" aesthetics of an environment somehow affect human productivity within that environment. In terms of designing habitats for space missions, we require empirical evidence to support specific contentions, otherwise detractors will regard aesthetics as a "sissy" issue. This is true for NASA and it is especially true for DoD missions. It is necessary to establish the relative importance of habitat aesthetics to the attainment of mission objectives in order to identify the trade-offs involved.

Outside Communications. For long-duration NASA missions, it is expected that capabilities will evolve early in the program for personal, two-way video calls to be made on a regular basis. It is generally recognized that periodic communication with family is beneficial in maintaining crew morale (e.g., distant-water fishing fleets, Salyut-6, etc.). The degree to which this will be possible on long-duration Air Force missions may be a matter of security. If communication is not allowed, behavioral consequences must be identified and anticipated.

Recreational Opportunities. Most of the available information on which we might base decisions regarding recreational issues is of an anecdotal nature. What is required is a systematic approach to the subject. Analogous earthbound conditions and previous experiences in space must be critically evaluated to identify appropriate recreational opportunities. We must not rely excessively on Skylab missions, however, when searching for lessons to be learned; the interests of future long-duration crews may differ signifi cantly from those of the past. That is, to satisfy the recreational needs of personnel on long-duration, routine missions, a radical and creative departure from previous efforts may be necessary. On one extreme, we may find it valuable to offer university extension courses leading to an advanced degree in applied sciences; at the other extreme what may be required is an element of friendly competition between crews. These suggestions are offered to illustrate the range of possible solutions that may be required to augment the traditional methods of recreation employed during short-term missions. Further research is required to identify acceptable and stimulating recreational opportunities for extended STS and long-duration missions.

Privacy and Personal Space. Much has been written about the need for privacy and personal space, most notably Helmreich's suggestions based upon evaluation of analogous conditions. What is yet required is a synthesis of these and other sources of data on the role played by these issues under conditions of isolation and confinement. While this concern is of paramount importance to long-duration missions, it is also important to short-duration missions (especially with larger crew sizes) and to survivable ground control facilities. Habitat design should acknowledge the dynamic and/or evolutionary nature of the system; what may be an appropriate design solution in the early stages of development may not be appropriate as the system matures. Further research is necessary to specifically define privacy and personal space guidelines for potential Air Force missions and to explore the extent to which habitats can be reconfigured internally to accommodate changing human requirements.

Waste Disposal and Management. Issues of waste disposal and management have been a particular crew concern since the first U.S. space mission. From the beginning, inadequate solutions have been developed for missions of short duration. Special attention must be directed toward the development of improved waste disposal systems for long-duration space missions. This should begin with the identification of the personal requirements and expectations of potential users in order to establish the effects, if any, on human performance resulting from inadequate waste disposal equipment.

Onboard Training, Simulation, and Task Preparation. During both extended and long-duration missions it is likely that additional onboard training will be required during the early stages of the mission for tasks to be performed during a later stage. This preparation may involve task simulation, review of procedures and specifications, focused discussion with ground control and/or crew personnel, or the physical inspection of equipment to be used during a task. In any event, it is likely that some form of onboard work station or learning center may be required to accommodate these preparation activities. Research is required to identify the types of preparation activities anticipated, and to develop equipment (e.g., flexible work station) to accommodate these requirements.

Behavioral and Physiological Requirements Associated with a Microgravity Environment. In addition to specific exercise, hygiene and food preparation requirements demanded by the absence of gravity, there are several related yet unresolved issues associated with a weightless condition. These include the design of consoles to effectively accommodate personnel in a neutral gravity position, the design of workbenches to hold small parts in place (in Skylab they used the exhaust filters), development of foot restraints, eating utensils, and storage/labeling systems, to name a few. Faulty solutions to these and other hardware issues have caused problems on previous U.S. space missions. These issues require resolution for human performance to be optimized on **future** missions. What is required is a systematic analysis of previous experience of living and working in space to identify the relevant issues and to develop effective solutions. Current NASA studies, proposed studies, and an evaluation of Spacelab missions may provide the needed data to satisfy these concerns.

SUMMARY

In the previous paragraphs we have focused on long-duration space missions in order to identify the habitability issues confronting the Air Force Space Command. **It must be understood that the issues discussed and the research requirements identified apply to long-duration missions, extended duration STS missions, and to lesser degrees, to human performance on short-duration missions and to the personnel staffing terrestrial space operations centers.**

Solutions to some of the problems associated with habitability issues are either available or in the process of development. Clearly, though, **additional research is required concerning each issue listed** in order to truly optimize human performance and productivity during operations proposed by the Air Force Space Command.

Motion Simulation with a G-Seat System: Sensory and Performance Mechanisms

Grant R. McMillan
Air Force Aerospace Medical Research Laboratory

William H. Levison
Bolt Beranek and Newman, Inc.

Edward A. Martin
USAF Aeronautical Systems Division

Abstract

Including whole-body motion in a flight simulator improves performance for a variety of tasks requiring a pilot to compensate for the effects of unexpected disturbances. A possible mechanism for this improvement is that whole-body motion provides high derivative vehicle state information which allows the pilot to generate more lead in responding to the external disturbances. In developing new motion simulation algorithms for an advanced g-cuing system we were, therefore, surprised to discover that an algorithm based on aircraft roll acceleration produced little or no performance improvement. On the other hand, algorithms based on roll position or roll velocity produced performance equivalent to whole-body motion. This paper describes the analysis and modeling being conducted at both the sensory system and manual control performance levels to explain the above results.

Introduction

The Air Force Aerospace Medical Research Laboratory and the Aeronautical Systems Division are jointly investigating motion and force cuing alternatives to whole-body motion. This paper summarizes the progress on an investigation of the capability of an advanced g-cuing system to provide rotational motion information to a pilot performing a flight control task. Human performance modeling is being conducted to explore hypotheses concerning the underlying sensory and performance mechanisms.

Roll-Axis Drive Algorithm Development and Testing

Two motion cuing devices were used: (1) the Advanced Low Cost G-Cuing System (ALCOGS), and (2) the Roll-Axis Tracking Simulator (RATS). The ALCOGS includes hydraulically-actuated seat pan, backrest, and seat belt elements mounted in an aircraft seat frame (Kleinwaks, 1980). In the studies reported here, the one-piece seat pan was the only active cuing element. The RATS is a whole-body, roll-axis motion device. The axis of rotation is through the buttocks of the subject.

The initial approach was to develop a means of driving the ALCOGS seat pan such that the pressures produced on the human buttocks matched those one would experience in the RATS. Using small force-sensing strain guages located under the ischial tuberosities of the buttocks, we measured the pressures produced by

sinusoidal roll motion in the RATS. A multiple regression performed on data collected over a range of amplitudes and frequencies suggested that buttocks pressures were a function of RATS roll angle and roll acceleration:

$$PSI_{Buttocks} = -.064\ \emptyset_{RATS} + .0041\ \ddot{\emptyset}_{RATS} \tag{1}$$

where PSI = pressure in lbs/in^2, $\emptyset$ = roll angle in deg, and $\ddot{\emptyset}$ = roll acceleration in deg/sec^2. For data collected under similar sinusoidal motion in the ALCOGS, buttocks pressures were a simple function of seat pan roll angle:

$$PSI_{Buttocks} = .081\ \emptyset_{ALCOGS} \tag{2}$$

Setting the equations equal to one another and solving for the ALCOGS seat pan angle (in deg) results in the following pressure matching algorithm:

$$\emptyset_{ALCOGS} = K\ (-.79\ \emptyset_{RATS} + .052\ \ddot{\emptyset}_{RATS}) \tag{3}$$

K values only up to 0.4 (40% of RATS pressures) were used to prevent the ALCOGS seat pan from striking its limits of travel.

The utility of this algorithm was evaluated by comparing human performance on a roll-axis tracking task under static (visual cues only) and g-seat motion conditions (visual and g-seat cues). The visual display consisted of an aircraft symbol and a dotted reference line, and subtended a 9 deg field-of-view. The task was to maintain zero roll angle (keep the symbol and reference aligned) in the presence of strong turbulence using a side-mounted, force-sensing control stick. The turbulence was generated as a sum of 13 sine waves. The roll dynamics were a first-order filter approximation to a heavily loaded fighter aircraft. Under motion conditions the seat pan of the ALCOGS was driven in roll using Equation 3. A group of six subjects was alternated between static and motion sessions, which consisted of four trials lasting 3 min each. The root-mean-squared (RMS) tracking error was displayed to the subjects after each trial.

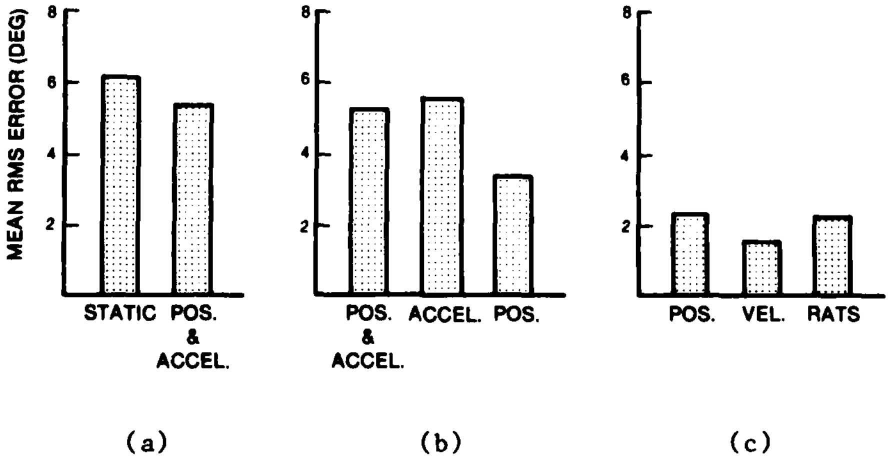


FIGURE 1. MEAN RMS TRACKING ERRORS UNDER STATIC AND MOTION CONDITIONS

The results are summarized in Figure 1(a), which shows mean RMS errors after the subjects reached asymptotic performance levels. The difference between static and motion errors was small, but statistically significant, $\underline{t}$ (5) = 6.08, $p < .01$. As shown in Figure 1(b), the next step split the position plus acceleration (pressure matching) algorithm into its two components. The pure position and pure acceleration algorithms were derived by setting the acceleration or position coefficients of Equation 3 to zero. Two of the six subjects were retained as controls on the original algorithm, and the remaining four subjects received the position and acceleration algorithms in a counter-balanced fashion. The data for the position and acceleration algorithms in Figure 1(b) do not represent asymptotic performance, but they do strongly suggest that the position component was providing much more useful information than the acceleration component. They also suggest that the acceleration component, when combined with the position component, partially masked the useful information.

This unexpected result suggested that an algorithm based on roll velocity might be of interest. In the velocity algorithm, ALCOGS seat pan angle was made proportional to simulated aircraft roll velocity:

$$\phi_{ALCOGS} = -.23\ \dot{\phi}_{Simulated\ Aircraft} \qquad (4)$$

where $\dot{\phi}$ = roll velocity in deg/sec. Investigation of the position algorithm also continued (shown here in terms of simulated aircraft roll angle):

$$\phi_{ALCOGS} = -.32\ \phi_{Simulated\ Aircraft} \qquad (5)$$

The results of a separate pilot study with the position and velocity algorithms, and with whole-body motion (RATS) are summarized in Figure 1(c). The RATS drive algorithm matched the roll angle of the simulated aircraft in a 1:1 fashion. In this study two groups of three subjects were trained to asymptote on one of the ALCOGS algorithms and were then trained to asymptote in the RATS. As shown in Figure 1(c), both algorithms produced performance equal to or better than whole-body motion.

Analysis and Modeling

To explore the performance enhancement observed with the ALCOGS position and velocity algorithms, human operator describing functions were calculated and plotted (Figure 2). The input and output signals for these describing functions are the visually displayed error and the control stick force, respectively. Compared to static tracking, both g-seat algorithms produced a substantial increase in low frequency gain, an increase in phase lead at higher measurement frequencies, and reduced remnant power at low and middle frequencies. This pattern of results is typical of comparisons between static performance and whole-body motion with a disturbance task (Levison and Junker, 1977). The slightly higher gain and phase lead seen with the velocity algorithm were reflected in the reduced RMS errors with this cuing scheme (Figure 1).

To investigate the mechanisms producing improved performance with the position and velocity algorithms, model analysis using the Optimal Control Model (OCM) is being conducted. The OCM (Kleinman, Baron, and Levison, 1971) is an information processing model developed within the framework of modern

control theory. It assumes that a well-trained and motivated human controller behaves optimally to achieve a specified performance criterion, subject to inherent perceptual-motor limitations. To use the model to predict performance on a given task one must specify: (1) the sources of information available to the human, and (2) model parameters which represent human perceptual-motor limitations. When using the model to "fit" existing data, one typically specifies certain of the information and perceptual-motor parameters and iteratively searches for values of the remaining parameter(s) which give a good model match.

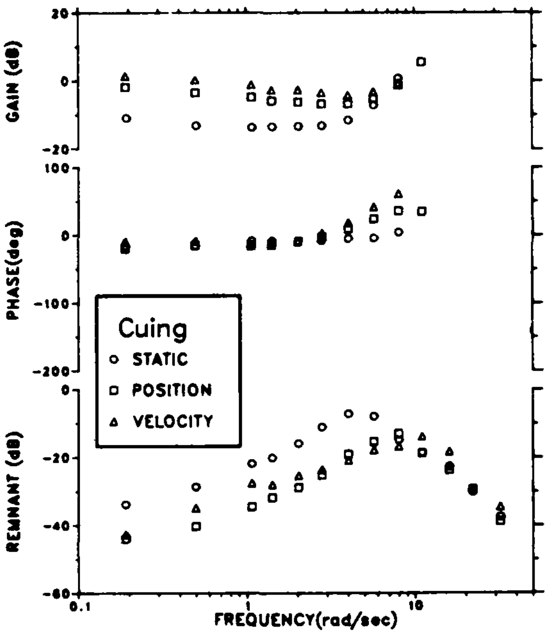


FIGURE 2. AVERAGE HUMAN OPERATOR DESCRIBING FUNCTIONS

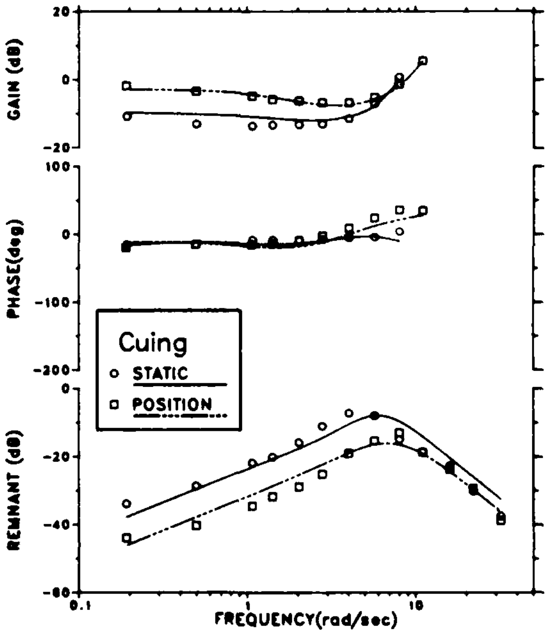


FIGURE 3. MODEL FITS TO HUMAN OPERATOR DESCRIBING FUNCTIONS

Initially, we determined the extent to which "nominal" parameter values (i.e., values typical of those obtained with model fits to previous data of this type) could describe the data shown in Figure 2. The model fit to the static results (Figure 3) was obtained entirely with nominal perceptual-motor parameter values: (1) -20 db observation noise/signal ratio which represents a sensory differential threshold (i.e., Weber fraction) of 0.10, (2) 0.2 sec time delay which represents peripheral and central processing time, (3) -50 db motor noise/signal ratio which reflects uncertainty in the pilot's knowledge of vehicle response, and (4) 0.1 sec motor time constant which may represent motor system bandwidth limitations. In addition, it was assumed that the information from the visual display included tracking error and error rate. Although a better fit could be obtained by "fine-tuning" the parameters, the nominal set produced an acceptable match.

The fit to the position algorithm data in Figure 3 was obtained by making two modifications. First, we assumed that error and error rate information were also available from the motion cues and that, in addition, the g-seat provided error acceleration information not available from the visual display. (It is well known that there are mechanoreceptors in the skin and

joints which provide 0th, 1st, and 2nd derivative information with respect to deformation of the skin or movement of the joints - Horch, Tuckett, and Burgess, 1977). Second, an observation noise/signal ratio of -27 db was assigned, based on an iterative search, to the information obtained from the g-seat (This represents better sensitivity to the seat cues). All other parameters were the same as for the static condition. As shown in Figure 3, this model fit captures the essence of the static-motion differences. The velocity data were fit (not shown) in a similar fashion using a noise/signal ratio of -30 db. This represents even better perception of information and is reflected in superior performance with this algorithm (Figure 1).

Attempts were made to fit the above g-seat data by manipulating other model parameters such as time delay. Although the data could be fit, unreasonable parameters for human processing (e.g., 0.007 seconds) were required. Thus, the model analysis suggests that the results can be parsimoniously accounted for by assuming that the human is more sensitive to motion cues, and that motion provides higher derivative tracking error information. Although this assumption is consistent with the model analysis of previous whole-body motion data (Levison and Junker, 1977), it clearly does not account for the poor performance with the pressure matching and pure acceleration algorithms. If one were to make the reasonable assumption that these algorithms also provided error acceleration information, the OCM would predict good performance.

To account for this anomaly, further OCM analyses are being conducted to evaluate alternative hypotheses. The first hypothesis assumes that, rather than providing separate channels of information related to the 0th, 1st, and 2nd derivatives of the signal driving the seat pan, the tactual system provides a composite cue containing a weighted sum of these derivatives. The composite cue produced by the acceleration algorithm would, therefore, contain the 2nd derivative of acceleration. This signal is likely to be so noisy that it masks the other information in the composite. A second hypothesis retains the assumption of separate channels of information, but will include the filter characteristics of the receptors believed to provide the various derivatives. It is possible that the receptor filter characteristics are interacting with the frequency content of the signals to produce differential filtering of the derivative information in the different algorithms.

References

Horch, K. W., Tuckett, R. P., & Burgess, P. R. (1977). A key to the classification of cutaneous mechanoreceptors. Journal of Investigative Dermatology, 69, 75-82.

Kleinman, D. L., Baron, S., & Levison, W. L. (1971). A control theoretic approach to manned-vehicle systems analysis. IEEE Transactions on Automatic Control, AC-16, 824-832.

Kleinwaks, J. M. (1980). Advanced low cost g-cuing system (AFHRL-TR-79-62). Brooks Air Force Base, TX: Air Force Human Resources Laboratory. (NTIS)

Levison, W. L., & Junker, A. M. (1977). A model for the pilot's use of motion cues in roll-axis tracking tasks (AMRL-TR-77-40). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory. (NTIS)

AD-P003 268

The Measurement of C-5A Aircrew Performance

Wayne L. Waag and David C. Hubbard

Operations Training Division
Air Force Human Resources Laboratory
Williams AFB AZ 85224

ABSTRACT

This paper describes efforts to develop and validate a performance measurement system (PMS) for C-5A aircrews. To date, a PMS has been developed and successfully integrated with one of the C-5A flight simulators at Altus AFB, OK. The system is currently being used as a routine part of the simulator training curriculum and is undergoing evaluation. This paper describes the capabilities of the simulator PMS and presents the results of the preliminary data gathered during the validation phase of the project.

Introduction

In the late 70's, the need for improved measurement capabilities in the operational training environment was realized and a major R&D program was initiated. In order to document the effectiveness and efficiency of the simulator training syllabus, it seemed desirable to quantitatively assess aircrew proficiency both in the simulator and the aircraft. The development of objective performance assessment capabilities for operational flight training systems was viewed as a two-phased process. The first phase would focus on the implementation of a measurement system within the flight simulation environment, while the second would focus on the aircraft. Two target applications were eventually selected: first, the development of a PMS for the C-5A transport aircraft, which is the focus of the present paper and second, the development of an air combat maneuvering PMS applicable to both the simulation environment and instrumented range facilities.

Description of C-5A PMS

A front-end analysis was initiated to define requirements for an objective PMS for members of the C-5A aircrew utilizing the existing C-5A simulator. Based upon the functional capabilities described in the final report (Swink et al, 1978), a contract was awarded for the development, fabrication, and integration of a measurement system for the C-5A flight simulator. The resulting system was installed on one of the simulators located at the C-5A Training Squadron at Altus AFB OK. Acceptance testing was completed in October 1982. Currently, the system is undergoing a two-year evaluation. Before describing the evaluation and some of the initial findings, a brief description of its functional capabilities is presented.

Mission Control. The C-5A PMS provides the potential for a precise definition of mission profiles whereby the proficiency of each aircrew member can be determined. The level of detail for the profile corresponds to that of a complete specification of student responses and aircraft/environmental factors (conditions, standards, etc.) to be addressed by the system. The

capability is also provided for the creation of new mission scenarios by instructor personnel within Military Airlift Command (MAC). The C-5A PMS has the capability to operate in either a preprogrammed or manual mode. In the preprogrammed mode, the PMS does not allow any alteration of the predefined mission; only monitoring and feedback controls are available. The manual intervention mode, however, provides the instructors with the ability to completely control the PMS mission. Predefined malfunctions may be activated or cleared at any time and additional malfunctions not contained in the original PMS scenario may be inserted. Of course, the instructors may allow the PMS scenario to operate as initially designed by the course developers.

Performance Monitoring, Assessment, and Analysis. The C-5A PMS monitors most switch positions, control settings, and instrument readings from the flight deck. It provides a real-time evaluation of proficiency for all crew members individually, as well as the coordination among aircrew members. Proficiency evaluations can be accomplished using the preprogrammed mission profiles. It requires that mission essential/critical flight tasks be included within the preprogramed profiles and that performance standards be precisely defined. The system provides a measurement capability for checklists, procedures, navigational profiles, and aircraft state parameters. It also has the capability to store, retrieve, and summarize all performance measures.

Performance Feedback and Displays. The PMS provides the following display and feedback capabilities: (a) mission sequence display--summary displays of the sequences of tasks; (b) route chart display--graphic background displays corresponding to departure, enroute, and approach plates; (c) checklist/procedure display--displays of predefined sequences of actions to be performed by crew members; (d) error alert display--message alerting instructor to crew errors as they occur in the predefined tasks; (e) proficiency assessment display--detailed alphanumeric displays relative to any specific predefined performance segment or task; (f) debriefing report--hard-copy containing objective performance data which the instructor may use for debriefing; and (g) help display--22 pages of on-line instruction on the use of the PMS.

To meet these functional capabilities, a "piggy-back" system configuration was required. In other words, the C-5A PMS is autonomous and independent of the host simulator in the sense of using its own processor and peripherals. All mission control, measurement computations, and display capabilities are accomplished with the PMS hardware. It is interfaced with the host simulator, so that it passively monitors all I/O signals (thereby obtaining all necessary data). It also has the capability to pass data actively to the host simulator for problem control and malfunction insertion, as necessary.

From this brief description, it should be apparent that the C-5 PMS provides a much wider range of capabilites than only performance assessment. It encompasses most functions which have been typically associated with the design of instructor/operator stations. As such, it can be considered a stand-alone instructional support system. Discussion now turns to the evaluation of the C-5A PMS and some of the initial findings.

Evaluations Plan for the C-5A PMS

The evaluation of the C-5A PMS will address the following areas: mission generation, operation, measurement, perceived operational utility, and

operational suitability.

Scenario Design Capabilities. The PMS scenario author has a number of difficult tasks. He must specify all relevant mission components in comprehensive detail. Flight checklists must be specified not only as to operation, but as to sequence. Navigational profile information must be included for the main flight path and any alternative flight paths among which the author might wish the students to choose. The PMS author is presently aided by several items. Different forms are provided to aid in constructing the various disk files that constitute a PMS scenario. A line text editor provides the author with the means of entering data from these forms into system and modifying it once it has been entered. The mission generation program provides a syntax and consistency check of the work. A consideration of this complex job leads to the following questions: Are the current aids adequate or do they need improvement? If they need improvement, how should they be improved? Are any additional aids needed? If so, what are they and how should they be integrated into the system?

System Operation. The PMS operation consists of two major types of operations: housekeeping and mission conduct. The housekeeping operations consist of the daily turn-on procedure, the confidence test, and the system shutdown. The mission conduct operations consist of commands to perform among other things the following: selection of PMS operating mode, performance monitoring, sign-on, real-time interaction with PMS monitoring, instructor station commands, and program termination. Of these two types of operations, it is the mission conduct operations that have the greatest impact upon the C-5A instructors. Another objective will be to assess these operations in order to identify the following: operations that are confusing or difficult to perform; operations that are unused or ineffective; operations that are not presently in the system that could prove helpful to instructors; suggestions for improving or replacing operations that are presently confusing or difficult to perform; and operations that instructors find helpful or effective.

Measurement Validation. At present, the PMS provides performance measures at five different levels. Level 1 measurement consists of assessment for individual tasks (e.g., lowering the landing gear before final approach). Each task is assigned a point value and tasks are scored according to a predefined algorithm. There are separate scores on each task for each crew member on each individual task. Scored tasks have been specified in advance by the scenario author along with the total possible point value for each task. The individual scores are then combined into three performance task groups for level 2 scores: checklist/procedures, monitorable parameters, and navigational profiles. Level 2 scores are computed by taking the total points earned under each group. If certain critical tasks are missed, the overall total is reduced by a predetermined criticality factor. Level 2 scores are obtained for each crew member for each of the three performance tasks groups. In addition, Level 2 scores are obtained for crew coordination in each of the three perforamance groups. If the scenario has two sessions, a separate set of Level 2 scores are computed from the Level 1 scores of the second session. Level 3 scores are computed for each crew member and crew coordination from the Level 2 performance measurement group scores. In computing Level 3 scores, each individual Level 2 score is first multiplied by a predetermined weight and then summed. The total score is then reduced by the appropriate

criticalty factors. The Level 3 scores are combined in a like manner to obtain an overall Level 4 score for each of the two scenario sessions. The two Level 4 scores are combined for a single Level 5 score for the entire mission.

As part of the test of the validity of the PMS assessment procedures, an analysis will be conducted to see how Level 2 and higher scores relate to instructor evaluations of crew performance. In addition, experienced flight instructor crews will be compared with inexperienced flight student crews on these PMS measures to determine if the scores can discriminate between experienced and inexperienced personnel. Since the assignment of weight and criticality factors involves a certain amount of subjectivity, different weight and criticality factors will be used in scoring. These new scores will then be tested for relation with instructor evaluation and discrimination ability. This evaluation should help to determine how much care is needed in assigning weight and criticality factors to the various levels of scores.

Operational Utility. The PMS is designed as a teaching aid for C-5A instructors. There are four basic teaching aids performed by the PMS: performance feedback to the student, objective mission assessment, mission monitoring for the instructor, and mission tailoring to student needs. Performance feedback is provided for the student in the form of a printed debriefing report. The report indicates which tasks were incorrectly performed, which flight parameters were out of bounds, etc. The report also contains the five levels of scoring as an objective assessment of the mission. During a simulated mission, the PMS provides the instructor a display which alerts him as soon as a student error is committed. The display also provides the means to monitor the progress of the mission, the course heading, the values of various parameters, etc. In the manual mode, the instructor has the freedom to alter the mission by inserting or clearing various simulated aircraft malfunctions, change selected simulated atmospheric conditions, suspend the session, skip over portions of the scenario, repeat portions of the scenario, etc. These instructor options provide the instructor with the ability to individualize the session to the students.

In order to determine the perceived utility of the PMS in the C-5A training program, the reaction of the instructors and students to the system is needed. In this regard, some questions include: Are the PMS displays and printouts conveying the necessary information or is there information not provided that would be useful? Is any of the information currently presented not needed? What PMS features are used most by the instructors? What additional features, if any, would the instructors like to have available to them?

Operational Suitability. Operational suitability factors pertain to how well a device meets accepted equipment serviceability requirements within its intended operating and maintenance environment. The major question is: How reliable is the PMS hardware in the operational environment? In order to answer this, a maintenance log and an instructor discrepancy log will be kept on the PMS during the entire test and evaluation period.

Preliminary Findings

The accomplishment of the above objectives required that the PMS be used in the various C-5A training courses. Three selected missions, Copilot

Missions 4, 5, & 6, have been designated as PMS missions. PMS scenarios have been developed for these missions and are currently used as part of the C-5A training curriculum at Altus AFB. As a result of this implementation several informal observations have been made relative to the operational utility of the PMS as perceived by the C-5A instructors.

Instructor Pilot Perceptions. Some of the initial observations included: (a) Navigational profile displays - These are considered the most popular feature among the IPs. As a result of the enthusiastic IP response, a graphics printer has been added to the PMS to provide copies of the ground track and glideslope displays for use in debriefings and between sessions; (b) Restart capability - Many IPs like the capability of restarting a navigational profile (e.g., an ILS approach) and having PMS reposition the simulator; (c) Multi-function keyboard - This is a source of difficulty for many IPs. As IPs gain experience in using the PMS, this problem may disappear; (d) Debriefing report - The order in which items appear on the report was not the order in which the IP would use them. Moreover, certain terms on the printout are not defined; and (e) Parameters monitoring - Most IPs do not like the parameters monitoring capability of PMS. Although they feel that maintenance of flight parameters is important, the quality of the basic simulator is such that it is impossible to maintain these within prescribed tolerances. In other words, the problem lies with the quality of the simulation and not the PMS.

Instructor Flight Engineer Perceptions. Some of the initial observations included: (a) Checklist/procedures monitoring - Most IFEs report that they like this capability, although the .8 sec sampling rate of the PMS is sometimes too slow to catch momentary switch positions. Many IFEs report that they like having PMS back them up when a student claims to have performed a step in a checklist when the IFE hasn't seen it; and (b) Automated malfunction insertion - Most IFEs like having the PMS insert the malfunctions. Without the PMS, the IFE follows a writen syllabus that indicates what and when malfunctions are to be entered. The IFE must manually insert the malfunction at the specified time by throwing a switch or turning a knob on the malfunction panel. A few malfunctions, such as fluctuation in oil pressure, require constant instructor input.

A note of caution should be exercised in that these findings represent some of the initial responses to the use of the PMS. These perceptions may change as instructors gain experience in the use of the system. Data collection for the evaluation will continue through Dec 84 so that all test objectives can be achieved. At this point, there seems little doubt that many of the capabilities of the PMS will be considered a useful adjunct to training. As such, the product of this development program will lead to a better specification of requirements for such capabilities in future simulator acquisitions.

References

Swink, J.R., Butler, E.A., Lankford, H.E., Miller, R.M., Watkins, H. and Waag, W.L. Definition of requirements for a performance measurement system for C-5A aircrew members. AFHRL-TR-78-54. Air Force Human Resources Laboratory, Williams AFB, AZ, Oct 78.

AD-P003 269

Aircrew Performance with Simulated Advanced Radar and FLIR Sensors in Single- and Two-Place Crew Stations

Roik L. Hockenberger
Hughes Aircraft Company

Abstract

The present study was directed at the use of synthetic aperture radar (SAR) and forward looking infrared (FLIR) sensor enhancements to quantify mission performance and aircrew workload effects. Both pilots and two-place crews were employed in a simulator based air-to-ground mission context. Results show significant performance and workload effects across the various system test conditions and between the single- and two-place crews. For example, performance scores revealed an average 15 percent improvement for the two-place crew, with an average 20 percent reduction in subjective workload rating.

Introduction

The effective utilization of advanced sensor capabilities in future fighter/attack weapon systems will be a key element of mission performance. The present study investigated a range of advanced synthetic aperture radar (SAR) and forward looking infrared (FLIR) sensor enhancements to identify and quantify performance and aircrew workload effects in an air-ground mission context. Both single-place and two-place crew configurations were employed to enable assessment of the potential advantages of a separation of functions and improved task loading in the two-place crew.

With manual SAR and FLIR sensor capabilities as a baseline, a variety of enhancements to those systems were considered that might offer superior tactical performance and reduced pilot workload. These enhancements included higher resolution SAR capability, automatic target detection, automatic classification capabilities, and automatic weapons handoff (Table 1). Aircrew tasks in the simulation included waypoint navigation, flight control, sensor operation/target detection, threat monitoring and response, handoff and delivery of WASP or Maverick weapons.

Method

A pre-post type repeated measures experimental design was selected wherein the manual SAR baseline (condition 1) was presented first, followed by the other six test SAR and FLIR conditions, and then repeated again in a final set of trials. This pre-post design enabled a measure of the training and experience gained by the subjects during the course of the test. With five subjects per crew size, a total of twenty observations were collected on each condition.

<u>Subjects.</u> Five pilots and five two-place crews were employed in the study. All were highly experienced Air Force TAC personnel.

Table 1. Sensor Enhancement Study Conditions

		SYSTEM CONDITIONS						
		BASELINE MANUAL SAR	MANUAL VHR SAR	SAR AUTO CLASSIFIER	LARGE AREA SAR/AUTO CLASSIFIER	MANUAL SAR/FLIR	MANUAL SAR/AUTO FLIR	AUTO SAR/ AUTO FLIR
CONDITION NUMBER		1	2	3	4	5	6	7
SAR DETECTION	MANUAL	X	X			X	X	
	AUTO			X	X			X
VHR SAR CLASSIFI-CATION	MANUAL		X					
	AUTO			X	X			
SAR POPUPS REQD		1	2	2	1	1	1	1
FLIR	MANUAL					X		
	AUTO						X	X
WEAPON	WASP	X	X	X	X			
	MAVERICK					X	X	X

Crew Station Simulators. The forward and aft cockpit crew stations used in the study were developed by Hughes Display Systems Laboratory as "general purpose" fighter/attack cockpits, composed of highly flexible displays and controls that can be rapidly configured in both hardware and software to meet research needs. In this way, it was possible to collect data in a highly cost-effective manner. The forward crew station is illustrated in Figure 1. Active displays included a wide field-of-view diffraction optics head-up display (HUD) and three multipurpose color CRT displays. Functions assigned to these displays included flight data, threat display, primary sensor display, and terrain following display. Threat lights and tones indicated threat type and status warnings. Conventional stick and throttle controls were used for flight control. The aft crew station, oriented toward sensor operation, provided a similar complement of displays and controls.

Performance measures were automatically collected during each trial. This allowed purposeful data analyses immediately following data collection without the need for time-consuming manual reformatting of the data. Following each trial, subjects were provided a workload rating scale form to score the workload experienced in each leg of the mission. The scale that was selected for this purpose was the Subjective Workload Assessment Technique -- SWAT (Reid, et al., 1981).

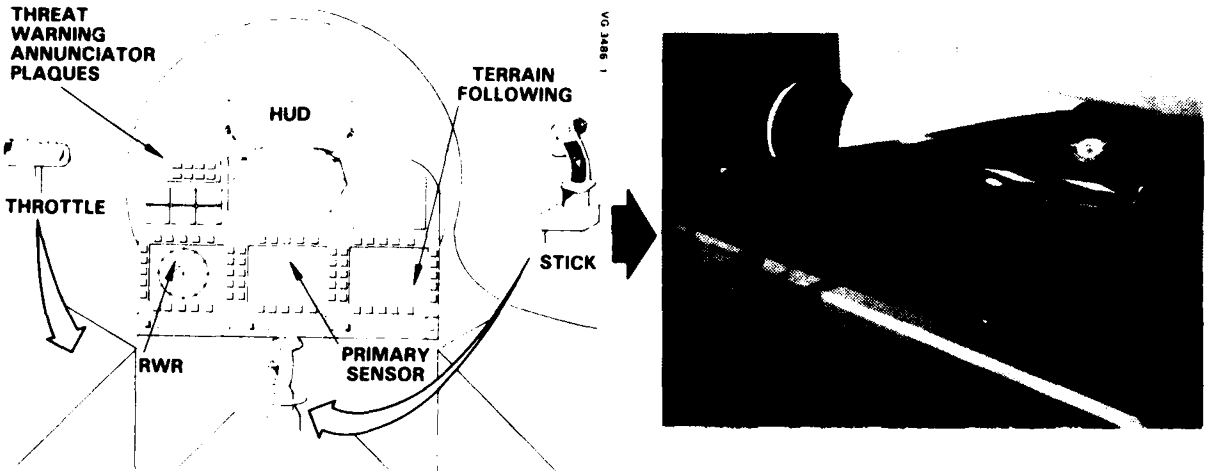


Figure 1. Forward Cockpit Crewstation.

Results

Performance. Performance results were collected by the mission tasks of target acquisition, weapon delivery, flight control, navigation, and threat response. Because of space limitations, only the target acquisition/weapon delivery measures are reported here. These measures are particularly salient because they show the cumulative effect of initial sensor operation and subsequent pilot launch decision performance.

Probability of a valid launch decision (Pvld) refers to the number of launches on valid targets plus the number of no-launch decisions on invalid targets. A plot across conditions on this measure, as provided in Figure 2, illustrates the crew size and sensor enhancement differences. The differences between sensor conditions were clearly evident for the pilots. For example, as enhancements to the Radar were added in conditions two, three and four, proportional increases in Pvld values were effected of 0.80, 0.90 and 0.95, respectively. Comparing the two enhanced FLIR conditions (six and seven) to condition five showed proportional increases in Pvld to 0.95 and 1.00, respectively. For the two-place crew size the sensor enhancement conditions did not show similar effects on Pvld measures. Equally superior values of 0.95 or better were found in all cases, compared to 0.70 for the manual SAR baseline.

Workload. Figure 3 illustrates the mean workload (SWAT) ratings for the weapon delivery mission leg. Crew size was a statistically reliable factor

(p <.0001), with the two-place crews rating much lower across conditions. Reliable differences were found among conditions (p<.016). The manual FLIR to Maverick weapon delivery condition scored reliably higher than comparison automatic FLIR conditions 6 and 7 (p <.05), as expected. The ordinal relationship of workload scores among conditions was consistent with expectations of the nature of sensor/weapons operations, and supported related performance results. This finding was encouraging relative to the utility of subjective workload techniques such as SWAT in controlled simulations of this kind.

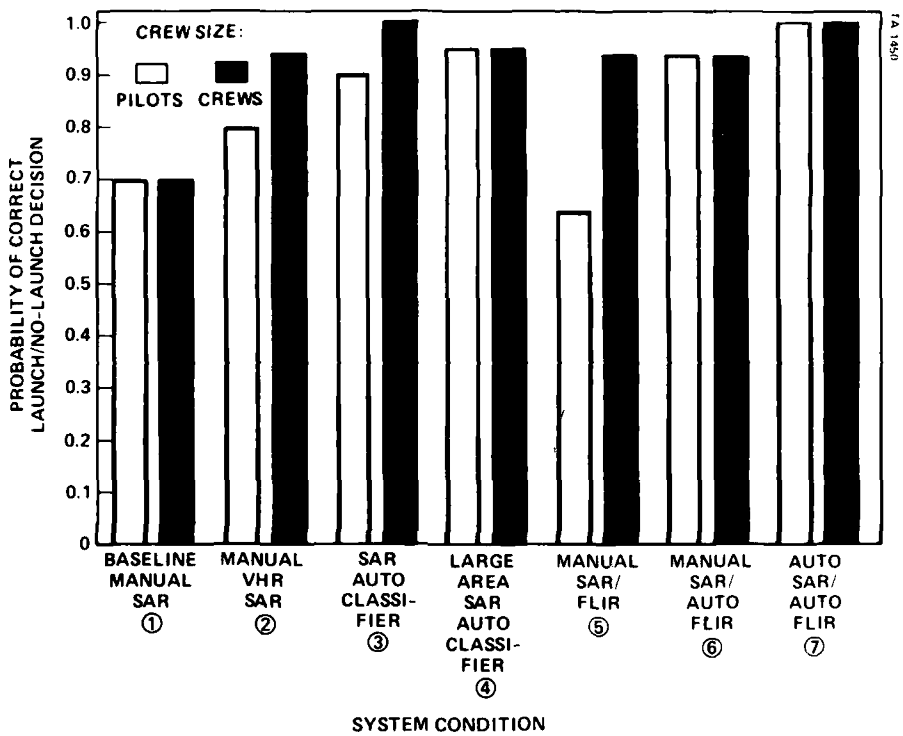


Figure 2. Probability of Valid Launch Decision (Pvld).

Conclusions

It is interesting to note that the performance differences between crew sizes diminished as more automation was incorporated into the enhancements. For example, in conditions four and seven, which represented the more automated SAR and FLIR conditions, the single-place pilot achieved equal performance to the two-place crews. Although pilot workload ratings never were as low as the two-place crews on a given condition, it was found that pilot workload on the more automated conditions tended to be rated lower than on the comparable baseline conditions. The present study has illustrated that the performance and workload impacts associated with a variety of sensor technology enhancements are incrementally discernible with a single place crew configuration. The two-place crew, on the other hand, is

able to achieve superior performance with both manual and automated systems, and with less workload differences. Future studies would be usefully directed at optimization of the single-place crew system through systematic application of cockpit automation enhancements.

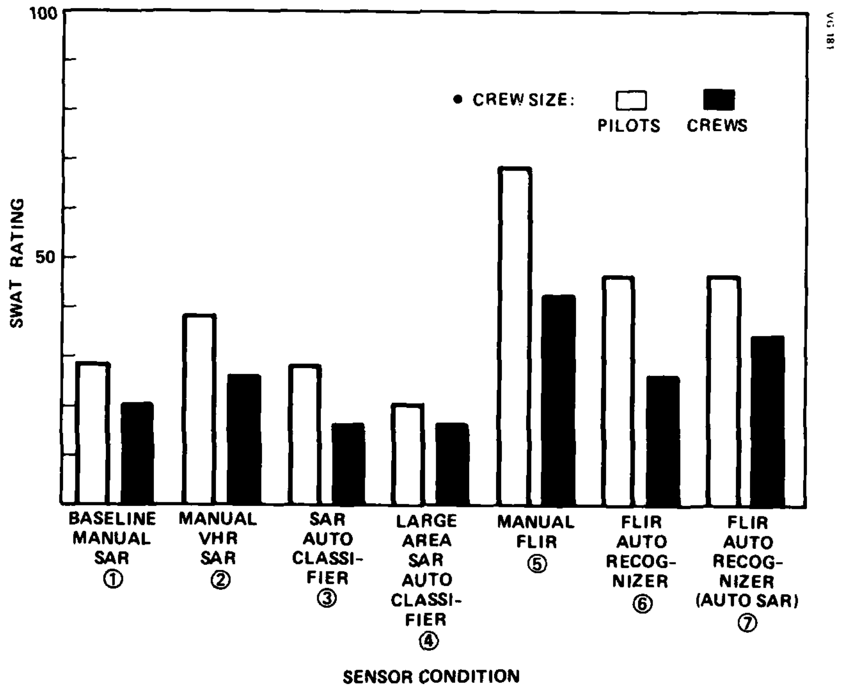


Figure 3. Workload Ratings for Weapon Delivery Mission Leg.

Reference

Reid, G.B., Shingledecker, C.A., Nygren, T.E., and Eggemeier, F.T. Development of Multidimensional Subjective Measures of Workload. Proceedings of the 1981 International Conference on Cybernetics and Society, October, 1981.

Acknowledgement

This research was sponsored by the Air Force Aeromedical Research Laboratory (AFAMRL) and Avionics Laboratory of AFWAL, under a contract entitled RAM/ACE (Radar Aided Mission/ Aircrew Capability Exploration), with McDonnell Aircraft Company as the prime contractor and Hughes Aircraft Company as a subcontractor making major contributions. The author gratefully acknowledges the assistance of Major Robert Eggleston of AFAMRL who offered many valuable suggestions relative to study design and data analyses. (Air Force Contract No. F33615-81-C-0501)

AD-P003 270

Health Risks Associated with Aircraft Model Type among U.S. Navy Pilots

Anne Hoiberg and Ralph G. Burr
Naval Health Research Center

Abstract

This study identified the ill health effects (hospitalizations) among U.S Navy male pilots (n = 22,245) who primarily flew one of eight aircraft model types. Comparisons of annual hospitalization rates computed for the July 1967 through December 1979 time period determined that trainer and helicopter pilots had the highest total rates while reconnaissance pilots had the lowest. Helicopter pilots had the highest hospitalization rates for ulcers and ear/hearing problems; rates for calculus of the kidney were the highest among older helicopter pilots. Older attack pilots had the highest rates for accidental injuries, musculoskeletal conditions, and infective/parasitic diseases. The highest circulatory disease rate was observed for older fighter pilots. An examination of the hospitalization rates of the other aircraft model types (electronic, patrol, cargo/transport, and reconnaissance) identified relatively few specific illnesses unique to each pilot group. Future research efforts will examine other factors in addition to aircraft model type that might account for the results of this study.

Introduction

In a recent study of the age-specific morbidity rates among Navy pilots (Hoiberg & Blood, 1983), comparisons of pilots' hospitalization rates with other Navy officer groups revealed an elevated risk among pilots for digestive diseases, accidental injuries, and musculoskeletal disorders as well as the specific conditions of ischemic heart disease, Hodgkin's Disease, cancer of the testis, cardiac arrhythmia, effects of heat, and bone rarefaction. Age differences were evidenced in that the youngest pilots had the highest rates for accidental injuries and disorders of tooth development and eruption while older pilots had the highest rates for circulatory diseases, particularly at ages 39-41. An important conclusion of that study was the need to determine whether or not the increased health risks for several aviation-related disorders could be attributed to the type of aircraft flown, especially for those older pilots with a high level of hours flown in a specific model. The purpose of the present study was to identify the ill health effects among Navy pilots who primarily flew one of eight different aircraft model types (fighter, attack, electronic, helicopter, patrol/antisubmarine, cargo/transport, reconnaissance, and trainer/all other aircraft).

Several researchers have addressed the issue of the pilot's physiological limitations to withstand levels of increased G forces of high-performance aircraft (Voge, 1980; Whinnery, 1982a; Whinnery, 1982b; Mills & Marks, 1982). Few studies, however, have identified the specific pathophysiologic effects

of repeated high G exposures other than to show that the cardiovascular system is the most sensitive to high G conditions and respiratory problems are considered to be second in importance. Similar to high-performance aircraft, other models entail their own special health-related consequences. Helicopter pilots, for example, would be expected to experience more lumbosacral pain, sciatica, trauma in joints and bone, visual acuity, hearing loss, and other effects of vibration than other pilots (Stave, 1979; Voge, 1981). Aircraft models modified with electronic counter measures might subject the pilot to an increased risk of experiencing a radiation-related disorder; pilots who fly sustained operations (e.g., cargo/transport) or who fly under adverse environmental conditions also experience greater demands on their physical and mental capabilities because of their increased workloads (Roscoe, 1978).

Method

Participants

Participants for this study were identified from the Individual Flight Activity Reporting System file which was provided to the Naval Health Research Center, San Diego, by the Naval Safety Center in Norfolk. Information on this file was compiled into individual pilot records ($\underline{n}$ = 22,245) consisting of a segment for each of 24 aircraft model types (e.g., A-4, F-14), which in turn were collapsed into the eight aforementioned aircraft model categories. On the basis of the highest total hours flown for one of the eight aircraft categories, each pilot was classified as primarily flying that specific type of model. For those pilots who had an equally high number of hours in two categories, they were assigned to the trainer/all other aircraft group; caution should be exercised in interpreting results for those pilots so assigned.

Procedure

To determine the age-specific health problems unique to pilots who primarily flew one type of aircraft, the number of hospitalizations for each diagnosis was tallied for the eight aircraft categories after dividing the subgroups into pilots 35 years of age and younger and pilots older than 35 years. All hospitalizations recorded from July 1967 through December 1979 were included in these tabulations for each diagnosis. Using mean populations at risk for each of the eight aircraft models by age, annual hospitalization rates per 10,000 strength were computed for each diagnosis for each pilot group. The χ^2 technique was performed to determine the level of statistical significance between a specific aircraft model group and all other pilots in both of the age-related subpopulations.

Results

Comparisons of Hospitalization Rates by Model among Younger Pilots

Comparisons of overall hospitalization rates between each model and the total rate of all pilots less than 36 years of age identified trainers/miscellaneous model pilots as having the highest total rate (628.2) of the eight groups. The only aircraft model group with a significantly lower total rate (255.3) was the group of reconnaissance pilots. Diagnostic categories with the highest rates for seven of the eight groups were digestive disorders and accidental injuries; cargo/transport pilots also had an elevated risk for genitourinary disorders.

Trainer/Miscellaneous. The highest hospitalization rates for this group were observed for digestive disorders, which primarily consisted of disorders of tooth development and eruption, and accidental injuries. Comparisons of cause codes for the accidental injuries revealed that only 6.3% of these hospitalizations were attributable to an on-duty, aviation-related mishap; 57.6% of all injury hospitalizations resulted from a sports-related or a non-military vehicular accident. Other significantly higher rates for this group were musculoskeletal disorders (primarily joint diseases), respiratory diseases, infective/parasitic diseases, skin disorders, symptoms and ill-defined conditions, and neoplasms. These pilots also had the highest rates for mental disorders, particularly alcoholism, although values were quite low.

Helicopter. Two categories and two specific diagnoses differentiated helicopter pilots as having significantly higher rates than others: diseases of the joint (although the rate was not as high as that for the trainer group), nervous system disorders, endocrine diseases, and ulcers. This group had the highest number of hospitalizations for the nervous system disorders of strabismus and ear/hearing problems.

Fighter. The third-ranked aircraft model in terms of total hospitalization rates was the fighter pilot group. The rates that differed significantly from other pilots included higher rates for symptomatic heart disease (with a value somewhat lower than cargo/transport pilots), symptoms and other ill-defined conditions, lipoma, and glaucoma. Fighter pilots were the only group observed as having any hospitalizations for glaucoma.

Electronic and Attack. Electronic pilots had the highest rates across groups for hernias and back disorders. Attack pilots differed from other groups in that they had a significantly higher rate for fractures, although their rates were lower than those for both trainer and fighter pilots. Their hospitalization rates for on-duty, aviation-related injuries tended to be the highest of all groups; such hospitalizations, however, accounted for only 13.3% of their total admission rate for injuries.

Cargo/Transport, Patrol, and Reconnaissance. The only significantly higher rates observed for cargo/transport pilots were genitourinary disorders and symptomatic heart disease. Patrol/antisubmarine pilots had significantly lower rates than other pilots for most diagnoses, and none of their rates was significantly higher. Reconnaissance pilots had so few hospitalizations (a total of 29) that rates for the majority of diagnoses could not be computed.

Comparisons of Hospitalization Rates by Model among Older Pilots

The only significant differences in total hospitalization rates across groups were the higher rate for helicopter pilots 36 years of age or older (401.1 per 10,000 strength) and the lower rate for reconnaissance pilots (123.1). Rates for the other groups ranged from 354.5 for attack pilots to 309.7 for fighter pilots. Overall differences in rates between younger and older pilots were minimal: 358.5 versus 338.4, respectively. Several shifts in rates, however, were observed in that rates for accidental injuries declined considerably between younger and older pilots for each of the eight aircraft model types, especially those for cargo/transport pilots. Other decreases in rates were observed for respiratory diseases, skin disorders, and infective diseases. Digestive disorder hospitalization rates showed considerable

fluctuation in that a decrease in rates for tooth development problems was observed as well as an overall increase in hospitalizations for hernias, intestinal disorders, and ulcers. The largest rate differential between younger and older pilots was noted by increases in circulatory disease rates, especially for helicopter, fighter, cargo/transport, electronic, and patrol pilot groups. Other increases in rates occurred for the categories of mental disorders (e.g., alcoholism) and neoplasms.

Helicopter. Older helicopter pilots had significantly higher hospitalization rates than other pilot groups for digestive disorders, ulcers, and calculus of the kidney and ureter. Also noted was the highest rate across groups for hearing loss hospitalizations among helicopter pilots.

Attack. Attack pilots' rates for musculoskeletal disorders (primarily back disorders), accidental injuries, infective diseases, and nervous system disorders were the highest of all pilot groups. Similar to younger attack pilots, the percentage of on-duty, aviation-related hospitalizations, although relatively low at 12.5%, was one of the highest across the eight groups. Eye diseases comprised 61.0% of all nervous system disorders.

Trainer/Miscellaneous and Patrol. Older pilots who were assigned to the trainer group did not have significantly higher rates than other pilots for any diagnostic category. The only significantly higher rate for patrol pilots than other pilots was noted for mental disorders.

Cargo/Transport and Electronic. Similar to helicopter pilots, cargo/transport pilots had a significantly higher rate than others for ulcers. The only diagnosis that even approached the level of significance among electronic pilots was chronic ischemic heart disease.

Fighter and Reconnaissance. Fighter pilots had the highest rate for circulatory diseases; a noteworthy shift in rates between younger and older fighter pilots was the decline in symptomatic heart disease hospitalizations and an increase in rates for chronic ischemic heart disease. Reconnaissance pilots had a total of four hospitalizations which precluded the computation of rates.

Discussion

Results of this study provided support for the general findings that increasing age was related to a decreasing accidental injury rate. Among younger pilots, injuries accounted for a high proportion of their total hospitalization rates although few of the injuries had an aviation-related cause code. Aviation-related deaths, on the other hand, accounted for a high percentage (62.2%) of all recorded deaths.

In addition to these general findings, other results identified several specific health risks postulated as related to piloting a particular aircraft model. Helicopter pilots had the highest hospitalization rates for ulcers, ear/hearing problems, and calculus of the kidney. These results lent considerable support for the research cited at the outset that described the adverse effects of high noise levels and vibration common to helicopters. While younger attack pilots had one of the highest rates for fractures and their older counterparts had the highest hospitalization rate for injuries, both subpopulations also had a somewhat higher on-duty, aviation-related injury

hospitalization rate than other pilots. With one of the highest mean number of hours flown, these pilots may have had a higher accident rate because of their higher level of exposure as well as the more stressful circumstances under which many of them flew. These occupational factors also may have contributed to their higher rates for back disorders and diseases of the eye. Fighter pilots differed from others because of their higher rate for circulatory diseases. Repeated exposure to high sustained G forces should be considered in interpreting these relatively high hospitalization rates. Interpretations of the results for the other aircraft model types were less clear-cut.

To conclude, the discussion suggests that the Navy pilot population is relatively unhealthy, a conclusion that should not be drawn from these results. According to previous research (Hoiberg & Blood, 1983), Navy pilots are an exceptionally healthy population when compared with the civilian community. Instead, the conclusion proffered is that subsequent research should be designed to identify the operational factors, personnel characteristics, and disease precursors related to the illnesses shown in this study to be at risk for pilots of such specific aircraft models as helicopter, attack, and fighter. Results of this proposed research, which currently is being initiated, will form the basis for the development of prevention and intervention health care programs. The benefits accrued from these programs will be to further enhance and protect the health status of all Navy pilots.

Acknowledgement

This research was supported by the Naval Medical Research and Development Command, Department of the Navy, under work unit MF58.528-01A-0001. The views presented in this paper are those of the authors. No endorsement by the Department of the Navy has been given or should be inferred. Requests for the two tables presenting the hospitalization rates for each pilot group may be mailed to the senior author.

References

Hoiberg, A., & Blood, C. Age-specific morbidity among Navy pilots. Aviation, Space, and Environmental Medicine, 1983, 54, 912-918.

Mills, F.J., & Marks, V. Human endocrine responses to acceleration stress. Aviation, Space, and Environmental Medicine, 1982, 52, 537-540.

Roscoe, A. H. Stress and workload in pilots. Aviation, Space, and Environmental Medicine, 1978, 49, 630-636.

Stave, A. M. The influence of low frequency vibration on pilot performance (as measured in a fixed base simulator). Ergonomics, 1979, 22, 823-835.

Voge, V. M. Acceleration forces on the human subject. Aviation, Space, and Environmental Medicine, 1980, 51, 970-980.

Voge, V. M. Role of pre-existing disease in the causation of naval aircraft mishaps. Aviation, Space, and Environmental Medicine, 1981, 52, 677-682.

Whinnery, J. E. Post-acceleration chaotic atrial rhythm. Aviation, Space, and Environmental Medicine, 1982, 53, 390-392. (a)

Whinnery, J. E. Acceleration-induced atrioventricular dissociation: hemodynamic consequences. Aviation, Space, and Environmental Medicine, 1982, 53, 432-434. (b)

AD-P003 271

Initial Validation of a Functional Age Profile for Aviators

Rolf Braune and Christopher D. Wickens
Aviation Research Laboratory-Department of Psychology
University of Illinois

Abstract

The initial validation of the concept of the "Functional Age Profile" for aviators is reported. This concept is based on objective performance parameters assessed by a computer-based information processing test battery. Preliminary results suggest that the "Functional Age Profile" may be an indicator of attentional resource capacity in flight simulator performance.

Introduction

Background

As the mean age of the population increases, concern has arisen regarding the impact of aging on human performance in a variety of complex systems. The present report examines the influence of aging on those information processing skills related to aviator performance. The reported work describes some of the results obtained during the second phase of ongoing research and development effort between the Aviation Research Laboratory of the University of Illinois and the Naval Aerospace Medical Research Laboratory, Pensacola, Florida (Braune and Wickens, 1983, 1984).

During Phase I of this research an information processing performance battery was developed that was designed to tap both aviation-relevant skills and age-related changes (Braune and Wickens, 1983). Factor analytic procedures conducted on the data from 60 subjects ranging in age from 20-60 revealed that the total set of 83 test variables could be reduced to five independent age-sensitive factors: (1) perceptual-motor speed; (2) perceptual-motor coordination; (3) focusing of attention; (4) spatial ability; and (5) field dependence/independence. The following representative variables for each factor were selected with high loadings on that factor: (1) Sternberg memory search RT with visual-spatial material, memory set size 4; (2) 2nd-order compensatory tracking, RMS error; (3) Dichotic Listening task, % Omissions; (4) Maze Tracing response time; and (5) Hidden Figures response time. Dual task measures were intentionally excluded because of the absence of any evidence that this ability changed with age in our sample. Using an individual's standardized performance scores on the five variables a "Functional Age Profile" can be developed. The purpose of the present study was to assess if variance in these age-sensitive measures was relevant to flight simulator performance.

Method

Subjects

Thirty males between the ages of 20 and 60 with 20/20 corrected vision and

normal hearing served as subjects. All subjects were rated pilots with instrument experience. Average flight experience was 800+ hours. Each subject was paid $3.00 per hour.

Tasks

The computer-based information processing performance battery was similar in format to the one used during Phase I. A detailed description of each individual test and the rationale for selecting it can be found in Braune and Wickens (1983). Table 1 shows a summary of all the tasks that were administered in single and dual task configurations. The flight task consisted of five 18 minute simulator flights through a series of instrument flight maneuvers shown in Figure 1. The flights were conducted in a Singer-Link General Aviation Trainer (GAT-2). The flight characteristics of the GAT-2 were modified to approximate those of a single-engine fixed-pitched propeller/fixed gear aircraft. Performance was recorded in terms of RMS-errors on the flight task and number-correct and reaction time on a secondary task simulating a communications task which was performed during certain segments of the flight. The total mission profile is shown in Figure 1.

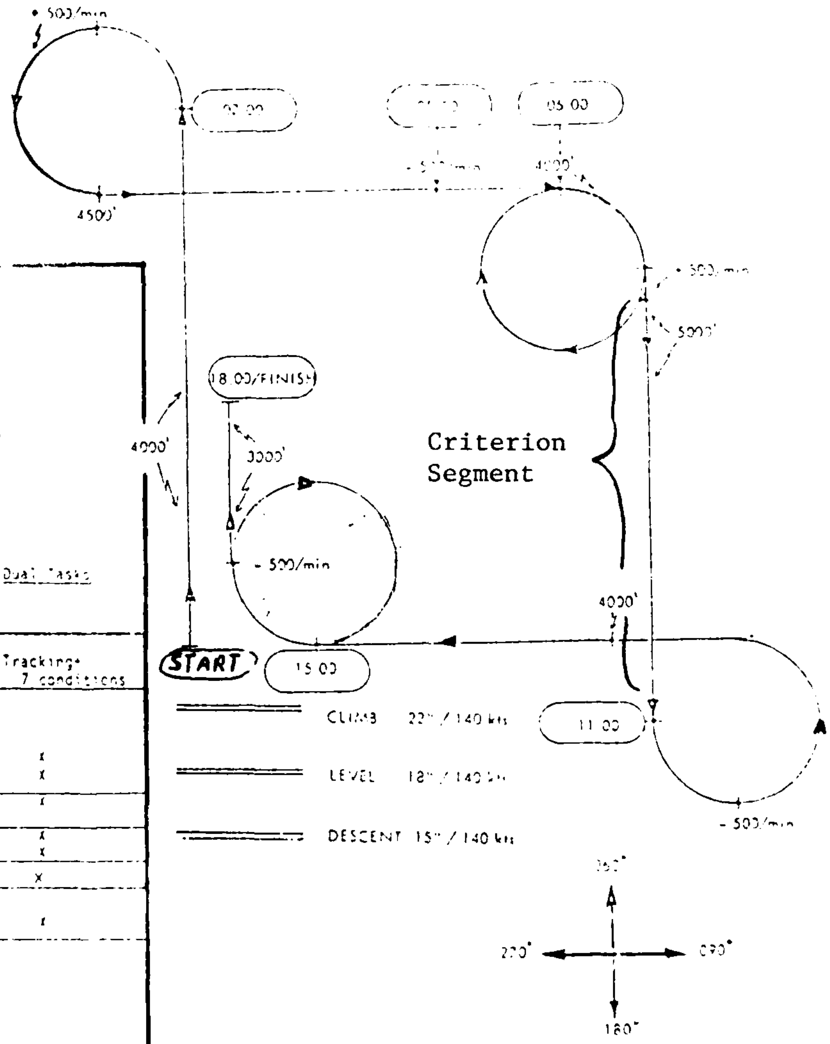


Figure 1: flight profile flown during the experiment

Table 1

Summary Table of the Information Processing Tasks Selected and Used in the Battery

Experimental Tasks

Tasks \ Modalities		Set-Size	Single Tasks: Stimulus	Single Tasks: Central Processing	Single Tasks: Response	Dual Tasks
Tracking	One-dimensional 2nd order comp.		V	S	M	Tracking+ 7 conditions
	Critical		V	-	M	
Sternberg (Memory Search Task)	1	2 4	V	V	M	x x
	2	2	A	V	M	x
	3	2 4	V	S	M	x x
	4	2	A	S	M	X
Delayed Digit Recall (2 dig. back)			A	V	M	x
Spatial	Card Rotation		V	S	M	
	Hidden Patterns		V	S	M	
	Maze Tracing		V	S	M	
Dichotic Listening			A/A	V	M	

V=visual A=auditory V=verbal S=spatial M=manual

Procedure

On the first day all subjects were administered the information processing performance battery equivalent to that used during Phase I (Braune and Wickens, 1983). On the second day each subject received a thorough simulator and mission check out. Flights 1-3 served as training flights. Flights 4 and 5 were the actual data collection flights. RMS-errors on five flight parameters were recorded by the computer, along with latency and accuracy of the secondary task. Total duration for Flights 1-5 combined was 3½ hours per subject. The average time between Day 1 and 2 was 4 weeks.

Results

The results reported represent a small subset of the complete analysis (see Braune and Wickens, 1984). Table 2 shows the intercorrelations between chronological age, flight experience, and performance on the Functional Age Battery. A relatively low relationship between chronological age and flight experience is indicated. The 2nd-order tracking task revealed the strongest correlations with flight experience.

Table 2 Intercorrelations between Chronological Age, Flight Experience, and Test Battery Performance

(p=.05, r=.32)

	1	2	3	4	5	6	7	8	9
1. Chronological Age	1.00								
2. Total Time	.29	1.00							
3. Simulator Time	.15	.41	1.00						
4. Instrument Time	.35	.81	.50	1.00					
5. 2nd-order Tracking (RMS)	.01	-.43	-.64	-.39	1.00				
6. Visual-Spatial (4) (RT)	.18	-.06	-.02	-.13	.25	1.00			
7. Hidden Figures (RT)	.36	-.06	-.34	-.08	.18	.35	1.00		
8. Maze Tracing (RT)	.08	.14	-.03	-.04	-.31	.26	.36	1.00	
9. Dichotic Listening (Om)	.48	-.04	-.01	-.09	.15	-.07	.21	.05	1.00

Table 3 Intercorrelations between Test Battery Performance and Flight Performance (Chronological Age and Flight Experience partialed out)

Criterion: Average Performance during Segment IV.

(p=.05, r=.32)

	1	2	3	4	5	6	7	8	[illegible]	[illegible]	[illegible]	[illegible]
1. 2nd-order Tracking (RMS)	1.00											
2. Visual-Spatial (4) (RT)	.29	1.00										
3. Hidden Figures (RT)	-.14	.32	1.00									
4. Maze Tracing (RT)	-.42	.23	.38	1.00								
5. Dichotic Listening (Om)	.12	-.28	-.01	-.04	1.00							
6. Altitude	.39	.34	-.08	-.21	.00	1.00						
7. Rate of Climb	.04	.34	-.17	.12	-.29	.20	1.00					
8. Heading	.17	.40	.29	.00	-.10	.14	.44	1.00				
9. Rate of Turn	.21	.14	-.10	-.14	.23	-.08	.58	.63	1.00			
10. Airspeed	-.08	-.29	.20	.38	.20	-.51	-.10	.21	.14	1.00		
11. Number Correct	-.03	-.65	-.32	-.39	.06	-.34	-.41	-.51	-.14	[illegible]	[illegible]	
12. Response Time	.23	.63	.25	.21	-.30	.16	.50	.25	[illegible]	[illegible]	[illegible]	[illegible]

For the purpose of criterion development the total 18 minute flight was broken down into 6 segments. Each segment was clearly defined by the maneuver that was performed and the workload demands imposed. The objective was to use that segment as a criterion which showed the least trade-off between primary and secondary task performance as indicated by the intercorrelations. Segment IV, a straight and level portion of the flight, indicated in Figure 1, best met this criterion, and so was selected. Table 3 shows the correlation table and correlations between the predictor variables and performance on the criterion segment. The correlations were run with the effects of age and flight experience partialled out. Those values within the box are the correlations between battery and flight performance.

It is apparent that the strongest relation with both flight and secondary task performance is shown by the spatial Sternberg Task. Furthermore correlations with the secondary task are generally higher than the primary. Multiple regression analysis was performed to determine the variance in flight and communications task performance that could be accounted for by the five battery items. The resulting equations and multiple R's are shown in Table 4. An average of 24% variance on the flight task and 46% on the communications task could be captured by the functional age battery. Correcting the results for shrinkage produced an average of 8% variance accounted for on the flight task and 34% on the communications task.

Table 4

Results of Multiple Regression Analysis Predicting Mission Performance from 'Functional Age Battery'

Criterion: Average Performance during Segment IV.

Criterion Variables	Equations	Multiple-R	Shrunken-R
Simulator Control			
Altitude	Y = -13.470 + .095(RMS) + .052(VS4) - .016(HF) - .001(MT) + .003(Om)	.53	.34
Rate of Climb	Y = 88.210 + .210(RMS) + .078(VS4) - .028(HF) + .004(MT) - .003(Om)	.47	.20
Heading	Y = -93.030 + .068(RMS) + .076(VS4) + .049(HF) - .002(MT) - .003(Om)	.50	.28
Rate of Turn	Y = 28.130 + .079(RMS) + .013(VS4) - .002(HF) - .0003(MT) - .002(Om)	.37	----
Airspeed	Y = 96.850 + .014(RMS) - .042(VS4) + .003(HF) + .003(MT) - .0001(Om)	.56	.39
Communications Task			
Number Correct	Y = 41.990 - .013(RMS) - .007(VS4) - .003(HF) - .0003(MT) + .00(Om)	.69	.60
Response Time	Y = 496.920 + .770(RMS) + .374(VS4) +.085(HF) + .005(MT) - .034(Om)	.67	.57

Discussion

The objective of the reported work was to validate the "Functional Age Profile" against operational performance as defined by a simulator flight task. The significant multiple R's suggest that age-dependent skills are relevant to

flight performance across the 20-60 year age span. The results, however, suggest that predictive power is stronger for secondary task performance than for primary task performance. Interpreting this result within the attentional resource framework it appears as if the "Functional Age Battery" could serve as an indicator of attentional capacity. However, more validation work in a variety of operational settings needs to be conducted before any final conclusions can be drawn.

Acknowledgement

This research was supported by contract N 00204-82-C-0113 from the Naval Aerospace Medical Research Laboratory, Pensacola, Florida. LCDR Dr. Jerry M. Owens was the technical monitor. The authors are indebted to Roger Marsh and Doug Heuer for the programming efforts.

References

Braune, R. and Wickens, C.D. Individual differences and age-related performance assessment in aviators. Part I: Battery development and assessment. University of Illinois Engineering Psychology Laboratory, Technical Report NAMRL-83-1/EPL-83-4, October 1983.

Braune, R. and Wickens, C.D. Individual differences and age-related performance assessment in aviators. Part II: Initial validation. University of Illinois Engineering Psychology Laboratory, Technical Report NAMRL-84-1/EPL-84-1, January 1984.

A COMPUTER SIMULATION TO PREDICT TARGET IDENTIFICATION LATENCIES

Dr. Valerie J. Gawron, Arvin/Calspan
Dr. K. Ronald Laughery, Jr., Arvin/Calspan
Dr. Charles C. Jorgensen, Oak Ridge National Laboratory
Dr. Joseph Polito, Pritsker & Associates, Inc.

ABSTRACT

As part of an ongoing program to develop computer models of human performance (MOPADS, Model of Operator Performance in Air Defense Systems), a computer model was developed to predict the target identification latencies of human observers. The model predicts these latencies as a function of one or more relevant independent variables. The values of these variables can be defined by the user (e.g., number of observers, observer's time on task) or be provided from an ongoing simulation of an air defense scenario (e.g., distance to target, target speed).

INTRODUCTION

As systems become more complex and more expensive, greater emphasis is being placed on evaluation of system performance during the design phase. In lieu of hardware models, more and more system designers are using computer models to evaluate alternate system designs. A key element in determining system performance from these models is a consideration of the performance of the human in the system. An accurate representation of the human's capabilities and limitations in a system model can greatly enhance the model's utility and applicability to the real world. Recognizing this fact, the Army has recently supported the development of a human performance model to be incorporated into a netted air defense system testbed. This effort, MOPADS, was completed last year.

MOPADS is written in the SAINT (Duket, et al., 1978) simulation language. SAINT uses activity networks to represent operator/machine systems. The components of the networks are "task nodes" which represent tasks or activities and "branches" which define the method by which the activities are sequenced.

In its current version, MOPADS incorporates human factors data in the form of moderator subroutines. These subroutines are called whenever the skill they describe is required by the task performance being simulated. In the subroutines, the time to complete and/or the probability of completing a task is calculated and the baseline nominal mean changed. The change can be in the form of a replacement (i.e., absolute moderator) or multiplication by a derived multiplicand (i.e., relative moderator). In either case, the focus is on the effects of various independent variables on a single operator's performance.

The method for developing the MOPADS moderator subroutine to calculate the time for an observer to identify a target is described below. A part of the subroutine itself is listed in the Results section.

METHOD

Skill Taxonomy

The performance of any task requires the operator to exercise one or more skills. Task performance can be moderated by calculating the effect of one or more

independent variables on the performance of the skills necessary to execute that task. Towards this end, a response-defined skill taxonomy was developed and is presented in Table 1. Recognition (or identification) is one of the skills identified in the taxonomy.

Table 1
MOPADS Skill Taxonomy

PROBABILITY ESTIMATION
TIME ESTIMATION
LONG-TERM MEMORY OF SENSORY INFORMATION
LONG-TERM MEMORY OF SYMBOLIC INFORMATION
SHORT-TERM MEMORY OF SENSORY INFORMATION
SHORT-TERM MEMORY OF SYMBOLIC INFORMATION
SYMBOLIC AND NUMERIC MANIPULATION
RECOGNITION
TIME SHARING
DETECTION
FINE MANIPULATION
GROSS MANIPULATION
GENERAL PHYSICAL EFFORT
REACTION TIME
TRACKING
TEAM COORDINATION

Literature Search

A computerized literature search (Laughery, 1981; 1982) was conducted to find human factors data relating independent variables of interest to the skills identified in the taxonomy. As a result, three categories of independent variables were identified:

1. environmental variables (e.g., ambient temperature);
2. operator variables (e.g., operator's time on task); and
3. task variables (e.g., the modality of the target to be detected).

Curve Fitting

Data identified during the literature search which related the time to identify a target to one or more independent variables of interest were then subjected to standard curve-fitting techniques to reduce the data to a single moderator equation. Seven types of equations were used (see Table 2). The form of the equation which accounted for the most variance (i.e., yielded the highest R^2) in the data was chosen for use in the recognition/identification model.

Table 2
Types of Equations

SIMPLE REGRESSION ($y=mx+b$)

SIMPLE REGRESSION - EXPONENTS ($y=x^m+b$)

SIMPLE REGRESSION - INVERSE EXPONENTS ($y=x^{-m}+b$)

SIMPLE REGRESSION - SQUARED EXPONENTS ($y=x^{m^2}+b$)

SIMPLE REGRESSION - INVERSE SQUARED EXPONENTS ($y=x^{-m^2}+b$)

MULTIPLE REGRESSION ($y=m_1x_1+m_2x_2$ . . . $+b$)

POLYNOMIAL REGRESSION ($y=mx^2+mx+b$)

Computer Logic

One of the goals from the very inception of the program was to maximize the match between the conditions being simulated and the conditions under which field data were collected. This required extensive use of branching in the code based on the current values of the environmental, operator, and task state variables. In the recognition/identification model, the major branches were defined by target modality (auditory, visual, or both) and observer-to-target position (ground-to-ground, air-to-ground, ground-to-air, or at-a-display).

In the case where multiple equations were relevant to the current state being simulated, the derived task completion times were combined in a weighted equation to produce a single derived time. The weight for each calculated time was the number of independent variables represented in the moderator equation from which that particular time value was calculated. Further, the difference between the values of the baseline mean and the time calculated from moderator equation N was divided by N. (The equations were ordered from largest to smallest effect on the dependent variable.) This approach is based on the law of diminishing returns which states that the inclusion of more predictor variables in a regression equation will increase the total amount of variance accounted for by the equation; however, each new variable added will account for a lesser portion of the total variance than any of the variables which preceded it.

RESULTS

A simplified portion of the recognition/identification model is presented below. T(NME) refers to the time for an observer to identify a target as calculated from equation number NME. In the following code, the effects of four independent variables are being calculated: slant range to target, target/background complexity, aircraft speed, and target type.

```
C
C**THIS IS AN AIR-TO-GROUND, VISUAL RECOGNITION TASK
C**XO(1)=SLANT RANGE TO TARGET IN NAUTICAL MILES
C**XFEET=SLANT RANGE TO TARGET IN FEET
C**XO(2)=TARGET/BACKGROUND COMPLEXITY
C**DATA FROM BIEDEMAN, GOMER, AND LEVINE, 1980,
C  PAGES 104, 158
C
      XFEET=X0(1)*6076.1
      NME=NME+1
      NIV(NME)=2
      T(NME)=(-13.812+0.001*XFEET+1.132*XO(2))/60.0
C
C**XO(3)=AIRCRAFT SPEED IN KNOTS
C**XFTSEC=AIRCRAFT SPEED IN FEET/SECOND
C**DATA FROM BIEDEMAN, GOMER, AND LEVINE, 1980,
C  PAGES 104, 158
C
    XFTSEC=X0(3)*1.688
    NME=NME+1
    NIV(NME)=1
    T(NME)=(-6.348+0.001*XFEET-0.010*XFTSEC)/60.0
C
```

```
C**XO(4)=TARGET TYPE
C**DATA FROM BIEDEMAN, GOMER, AND LEVINE, 1980,
C  PAGES 104, 158
C
C**TARGET IS A TANK
C
     IF (XO(4).EQ.11.0)THEN
       NME=NME+1
       NIV(NME)=2
       T(NME)=(-8.990+0.001*XFEET)/60.0
     ENDIF
C
C**TARGET IS A HALF-TRACK
C
     IF (XO(4).EQ.17.0)THEN
       NME=NME+1
       NIV(NME)=2
       T(NME)=(-12.57+0.001*XFEET)/60.0
     ENDIF
C
C**TARGET IS A TRUCK
       IF (XO(4).EQ.15.0) THEN
       NME=NME+1
       NIV(NME)=2
       T(NME)=(-13.220+0.001*XFEET)/60.0
     ENDIF
```

The end product is a single value for time to identify a given target under the conditions specified. To demonstrate the output, the model was run using all four variables under two sets of conditions. The results from these runs are listed in Table 3.

Table 3
Time to Identify a Target on a Display Derived
From The MOPADS Recognition Moderator Function

INPUT	CONDITION SET 1	CONDITION SET 2
Slant range to target	5 nautical miles	5 nautical miles
Target/background complexity	1	1
Aircraft speed	220 knots	220 knots
Target type	tank	truck
OUTPUT		
time to identify	19.8 sec	18.0 sec

DISCUSSION

Even the most technologically sophisticated systems are limited by the capabilities of the human operator(s) in these systems even if the human's role is only to activate the system. In many cases, simple changes to system interfaces and/or procedures in the design phase could have drastically reduced operational difficulties associated with the final product. Computer simulation during the early stages of a system's design can identify operational problems to come. Simulations which include models of human performance greatly increase a system model's utility. The recognition model described above can be easily incorporated into simulations of air defense and air traffic control systems. Further, each skill listed in Table 1 has an associated moderator function and the model can be expanded to include more data as such data become available. The human factors model is designed to stand alone from the MOPADS software and can be used in any setting that conforms to a well defined data interface. For the interested reader, the MOPADS Year 1 Report is available upon request (Polito and Laughery, 1982). Other moderator function subroutines are presented in Laughery and Gawron (1983).

ACKNOWLEDGMENT

This work is supported by the U.S. Army Research Insititute (contract number MDA903-81-C-AA06). Pritsker & Associates, Inc. is the prime contractor; ARVIN/CALSPAN the subcontractor. The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

REFERENCES

Biedeman, L.R., Gomer, F.E., and Levine, S.H. Dynamic target acquisition: Empirical models of operator performance. St. Louis, MO.: McDonnell Douglas Astronautics Company, AFOSR-TR-80-1177, August 1980.

Duket, S.D., et al. Documentation for the SAINT simulation program. Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory, AMRL-TR-77-63, June 1978.

Laughery, K.R. A data base for quantitative human performance modeling. Buffalo, NY: ARVIN/CALSPAN Corporation, MOPADS Vol. 5.2 DI, January 1982.

Laughery, K.R. A summary of the literature on quantitative human performance models. Buffalo, NY: ARVIN/CALSPAN Corporation, MOPADS, Volume 5.1 DF, December 1981.

Laughery, K.R. and Gawron, V.J. MOPADS skill moderator function subroutines. Buffalo, NY: ARVIN/CALSPAN Corporation, MOPADS Volume 5.10 DF, June 1983.

Polito, J. and Laughery, K.R. MOPADS technical report year 1. Albuquerque, NM: Pritsker and Associates, Inc., July 1982.

The design of manually operated controls for a six-degree-of-freedom groundborne "walking" vehicle: Control strategies and stereotypes

Dennis B. Beringer
The University of Wisconsin at Madison

ABSTRACT

The problems posed by control of "walking" six-degree-of-freedom vehicles vary with specific task requirements. Different strategies are required for cruising "supervisory" control and precision foot placement. These, along with anthropometric and stereotypic considerations, directly affect the design of the manually operated controls and the degree to which multiple-axis/multiple-task integration can be achieved within a single controller.

INTRODUCTION

Many studies have been done on the manual control of vehicles (see McRuer & Weir, 1969, for a review), some having six degrees-of-freedom of motion. A similar body of knowledge has been accumulating in the control of remote manipulators and "man amplifiers" (exoskeleton devices). The latter can be termed "anthropomorphic" devices as they mimic motions of the human body or appendages to some degree. Manipulators have generally been either fixed- base or transported by a separate propulsion system. The current case of interest is where the tail wags the dog; the end effector of the manipulator is placed against the ground to provide both support and propulsion for the entire system. The manipulator limb essentially grasps the ground and "moves" the earth relative to the "base" or shoulder end of the limb. This type of locomotion offers an alternative for terrain that cannot be traversed by wheeled or tracked vehicles (McGhee, 1984).

PROBLEM

Specific Vehicle Configuration

The vehicle currently under construction at the Ohio State University is a six-legged hydraulically actuated device. The legs are of mammalian rather than insect design for weight-bearing reasons and each leg has, within kinematic limits, three degrees of freedom (df). The vehicle body has (again within limits) six df. Although this may initially appear to be a 24-df problem, this is not precisely the case. Changes in body position/attitude are achieved by coordinated manipulation of leg positions so that these df's are not actually independent. One increases vehicle height, for example, by extending all legs equally. This vehicle was intended as a full-scale concept demonstration of microprocessor-controlled limb coordination. Manual control considerations have been, to some degree, after-the-fact as in many applied human factors problems.

Task Requirements

This vehicle will be required to perform tasks at the extremes of the conceivable operational range. At the most basic level it must be capable of individual foot placements for negotiating obstacles in rough terrain. This requires that each leg be controllable separately as well as body position relative to the foot placements. Present strategies call for alternately placing pairs of feet and shifting the body to maintain the vehicle within margins of static stability (McGhee, 1984). This is very much a manipulation/action-oriented activity and requires a high degree of operator control authority. The operator must select a foot placement that will (1) support the vehicle (terrain structure), (2) maintain static stability of the vehicle (support geometry), (3) be within the kinematic limits of the leg (reach envelope), and then (4) place the foot at that point. This requires the operator to provide inter-limb coordination and can thus be termed an "uncoordinated" system.

At the opposite extreme the vehicle is expected to be capable of traversing relatively level and obstructionless terrain with a minimum of operator intervention. The highest level of supervisory control, and lowest control authority, would have the operator specify only a positional goal to

which the system should travel. Means of goal accomplishment would be left to the vehicle systems. Although such an approach has been envisioned for future "autonomous" vehicle operation, present requirements are a step lower in the hierarchy; continuous control of direction, velocity, heading, and height of vehicle. Individual leg motions are not controlled by the operator but rather by a "coordination computer". Body roll, as an example, can be achieved by a single control input that causes legs on one side to flex and those on the other to extend. This can be termed a "fully coordinated" system.

DESIGN IMPLEMENTATION

Control Strategies

Precision footing. The case of individual foot placement (precision footing) is certainly accomplished better by resolved motion control (coordination of foot movement across leg joints or "intra-leg" coordination; Whitney, 1969) than by individual control of joint motions. [Such coordination software, in combination with force sensor information from the feet, has been demonstrated successfully in an earlier vehicle developed at OSU (McGhee, 1984).] It is also desirable to exercise control through a single integrated controller than through multiple one-dimensional controllers (Crawford, 1964) as are seen on numerous hydraulically actuated construction vehicles (McGhee, 1984).

The most direct approach appears to be a zero-order (position) 3-df control (position-following, force-feedback control; also called master-slave; Johnson & Corliss, 1971). Several factors argue against this in the present application. First, this type of system was used in the General Electric Quadruped Transporter (McGhee, 1984), a hydraulically actuated walking machine, and produced high levels of cognitive and physical fatigue in operators within relatively short time periods. Second, the operator's controls will be limited to one for each hand, requiring cycling through three legs with a single controller. A position- following control would encounter positional discontinuity during transition between legs, requiring a position reset for each. This would increase cycle times and undoubtedly cause consternation if not discomfort for the operator. This suggests a first-order (rate/velocity) control for x,y foot positioning. Foot lift/drop can be semiautomated using proximity sensing to execute final touchdown (available on current vehicle).

Cruise mode. Cruising at "high speed" (5-8 mph), the operator should need only to specify the resultant vehicle motion as leg motions are fully coordinated (ability already demonstrated on previous vehicle at OSU). A 4-df controller would be desirable as it would allow single-hand control of travel direction and velocity, heading (independently of direction of travel), and altitude. Pitch and roll are much less likely to require continuous control and thus can be trimmed by discrete means (4-position trim switch). A first-order (velocity) system seems preferrable as the operator then controls directly the indices of desired performance. This parallels the performance control system described by Roscoe and Bergman (1980) whereby direct control of aircraft attitude was traded for control of climb and turn rates, decreasing task performance error and operator workload but also decreasing control authority.

Another approach to cruise control has been suggested by Vertut (1983) that has considerable appeal. He has suggested that the operator, using a 3-df controller, set primitives of travel direction and path curvature (turning center) in the absence of vehicle velocity. One then applies along-path velocity after the path is determined. This has been quite useful for close maneuvering (obstacle negotiation) and is being examined for use in nuclear reactor inspection and maintenance vehicles. It effectively separates trajectory control from velocity control. Vertut has suggested an explicit controller design (physical device) discussed in a subsequent section.

Although precision footing and cruise control have been discussed separately, the ultimate goal is to define a physical device (control) that will serve in either mode of operation. Cruise control was the first problem to be considered and will receive more attention here. Precision footing is still being examined as of this writing.

Control Axes Assignments

Four factors are likely to influence the pairing of control motions with vehicle responses; population stereotypes, dynamic anthropometry, task frequencies, and interaction of control motion and vehicle motion. Only the first three will be addressed here as vehicle dynamics, though suspected, are not yet known fully.

Population stereotypes. It is generally acknowledged that human operators can and will adapt to a variety of control/response relationships that are nonpreferred, exhibiting accuracy close to that

achieved with preferred arrangements. The nonpreferred arrangement will often, however, produce longer response times, require longer training, and be subject to habit regression under stress (see Loveless, 1962, for a review). Although one might be tempted to apply single-axis stereotypes to a multi-axis controller, the data actually collected in this area have not been specifically related to more than two-axis control. It is suspected that interactions in the form of tradeoffs are likely to occur when multiple axes are available simultaneously, particularly when any ambiguity of assignments is possible. Thus it was decided to attempt a design given all available facts and perform "... a confirmatory check upon the actual design ..." (Loveless, 1962; pg. 381) regarding stereotypes.

Dynamic anthropometry. The hand is capable of three df of motion at the wrist: Ulnar/radial deviation (wrist pitch), flexion/extension, and rotation (see Bazar, 1978 and Garrett, 1971, for limits of motion). One can thus secure the operator's forearm and still have three df of continuous control that are less subject to disruption by vehicle accelerations. This approach was used by Bauerschmidt and Besco (1962) and produced a 3-df controller for manned space flight having rotational axes located in the wrist. The present design is an outgrowth of this approach and varies on only two points: wrist-pitch pivot point is located in the control grip (a mechanical consideration) and fore/aft sliding motion has been added. Wrist pitch is normally locked out until a trigger is depressed. A similar system with 6 df has been proposed by Roscoe, Hull, Simon, & Corl (1981).

Synthesis. Stereotypic responses favor control axes assignments that are direct analogs; fore/aft sliding motion of control implies fore/aft system motion. Given the available control motions, assignment by stereotypes suggests the following (by control motion/response motion): fore-aft/fore-aft, wrist- pitch/altitude, wrist-flexion/heading change, and wrist-rotation/lateral motion. High-frequency control activity should be paired with motions that are least difficult and/or fatiguing, however, wherever there is an ambiguous assignment (in particular wrist flexion versus wrist rotation). Fixed-base simulation of the system has indicated that heading-change commands are much more frequent in cruise mode than lateral motion commands. Inasmuch as repeated wrist flexion can lead to tenosynovitis (Bazar, 1978), it seems appropriate to exchange these assignments; wrist rotation becomes heading change and wrist flexion becomes lateral velocity. The arrangement is depicted in Figure 1. A conventional 3-df joystick (3rd df = rotation about vertical axis) is being used as a comparison (altitude is set by a separate control in this case).

The physical axial assignments and design suggested by Vertut for his controller are shown in Figure 2. This arrangement allows the operator to see where the turning point/center is physically located on the ground. An imaginary extension of the top horizontal bar of the yoke intersects the ground plane at the turning center (when the terrain is planar and the controller base is parallel to the terrain). Experimental performance comparisons are currently being conducted to assess relative training requirements and performance capabilities.

Confirmation of stereotypes. This allocation of axes may be seen as a slight conflict with stereotype responses if one accepts that control motion should be analogous to vehicle motion. In order to assess multiple-axis assignment preferences, 19 subjects ranging in age from 23 to 51 years (mean = 26.3, median = 29) (13 male, 6 female) were presented with a figure depicting the six df of motion of a rectangular solid (the vehicle). Each was then shown specific hand/arm motions related to controller operation and asked to assign each motion to the expected vehicle response motion. Two sets of motions, Series 1 having a rotational center in the operator's wrist and Series 2 having a rotational center beyond the operator's hand limits in two axes (conventional joystick), were presented (see Figure 3).

Single majority patterns emerged for each series, accounting for 47% and 42% of the sample respectively. In each case, vehicle motion was expected to be a direct analog of control motion. An additional 27% and 47% respectively transformed a rotary motion into linear vehicle response in the same "direction" as the rotation. The remaining individuals rotated responses 90 degrees in the plane of control motion (linear fore/aft or radial fore/aft = pitch; obtained from, of course, pilots).

The analogous assignments of particular interest, wrist flexion/heading change and wrist rotation/lateral motion, were chosen by only 10.5% as a first-choice assignment. This would probably have been much higher had vehicle roll not been allowed as a response option. Eleven subjects were, however, asked to give second-choice assignments and 36.4% did choose this arrangement then. The interchanged assignments (flexion/lateral motion, rotation/heading change) selected a priori were preferred, as a pair, by only 9% as a second choice. Although none chose both pairings as a first-choice pattern, 22% did choose the former and 16% chose the latter singly.

Although this assessment provided some insight into preferred pairings with unconstrained vehicle motions, it did not specifically address the "goodness" of the specific controller in the

constrained case (only 4 df of motion controlled continuously; pitch and roll trimmed discretely). A second examinatin was conducted where subjects manipulated the prototype controller and indicated preferred pairings between control motions and the now constrained class of vehicle motions.

Ten subjects, five male (ages 22 to 52 years; mean = 34.2, median = 32) and five female (ages 23 to 30; mean = 28.4, median = 30), were asked to perform two assignment tasks. In the first, subjects chose a single axis of control motion and rank ordered the four allowable vehicle response motions; this was done for each of the four control-motion axes (comparable to Set 1 of Figure 3). With each ranking the subject indicated whether the arrangement was "acceptable", "marginal", or "not acceptable". In the second task, each subject was asked to chose a single set of control-motion/vehicle-response pairings (a 1:1 mapping) that they would prefer for vehicle operation.

Kappa (Fleiss, 1973) was computed for each of the sets of data to assess the strength of agreement between the a priori design decisions for axes assignments and subject preferences (kappa (k) is similar to a correlation coefficient, ranging from -1 to +1). In the first case, unidimensional rankings (first choices only), k = .73 [z = 8.36, p<.0000001]. There was complete agreement on fore/aft and left/right dimensions, 90% agreement on up/down, but only 30% agreement on rotation (70% chose wrist rotation to represent left/right motion). This was the expected equivocation of axes, but it occurred for wrist rotation only and not wrist flexion as would be expected.

Computation of k for the second case (1:1 mapping) gave k = .93 [z = 10.22, p<.0000001]. In this case complete agreement was obtained, again, for fore/aft and up/down assignments. Only one individual exhibited reversal of left/right and rotational assignment (90% agreement). Thus 90% of the sample, when forced to choose a multidimensional solution to the problem, agreed with the assignments chosen in the design process, the only deviation being the expected one.

SUMMARY AND PROSPECTUS

It appears both possible and desirable to design a single integrated controller for manual control of an anthropometric "walking" vehicle, particularly when intra- and inter-limb coordination can be performed by computer components of the system. Without this aiding the rapid coordinated movements required for effective control would either not be possible or would impose very high workloads on the operator, limiting both the range and duration of activities. Problems posed by control/response stereotypes and dynamic anthropometry seem easily surmountable in most instances. It is clear that individuals make tradeoffs in multidimensional control/response assignments that do not reflect directly their unidimensional preferences. This reinforces Loveless' admonition to verify designs and provides an indication that some assignments may take priority over others. The challenge remaining is to deal effectively with the problem of precision footing, an area requiring both integrated controls and displays for successful vehicle operation. With this accomplished a wide range of applications may be opened for this class of vehicles, ranging from construction and timber harvesting to mobility aids for the handicapped. Successful applications, however, will ultimately depend upon the ease with which the tail can wag the dog.

ACKNOWLEDGEMENTS

This research is being conducted at the Human Factors Research Laboratory, University of Wisconsin - Madison, with contractual support from the System Sciences Division, Defense Advanced Research Projects Agency, Contract Number MDA 903-82-C-0028. The author wishes to thank Naomi Swanson and George Gruetzmacher for their assistance in data collection and figure preparation.

REFERENCES

Bauerschmidt, D.K. & Besco, R.O. Human engineering criteria for manned space flight: Minimum manual systems. Wright-Patterson AFB, OH: Aerospace Medical Research Laboratories, Tech. Documentary Report AMRL-TDR-62-87, 1962, 23-24.

Bazar, A.R. Ergonomics of hand tools. Hazard Prevention, November/December 1978, 10-13.

Crawford, B.M. Joy stick vs. multiple levers for remote manipulator control. Human Factors, 1964, 6(1), 39-48.

Garrett, J.W. The adult human hand: Some anthropometric and biomechanical considerations.

Human Factors, 1971, 13(2), 117-131.

Fleiss, J.L. Statistical methods for rates and proportions. New York: John Wiley and Sons, 1972, 144-147.

Johnsen, E.G. & Corliss, W.R. Human factors applications in teleoperator design and operation. New York: Wiley, 1971.

Loveless, N.E. direction-of-motion stereotypes: A review. Ergonomics, 1962, 5(2), 357-383.

McGhee, R.B. Vehicular legged locomotion. In G.N. Sardis (Ed.), Advances in Automation and Robotics. Jai Press, Inc., 1984, (in press).

McRuer, D. & Wier, D.H. Theory of manual vehicle control. Ergonomics, 1969, 12(4), 599-633.

Roscoe, S.N. & Bergman, C.A. Flight performance control. In S.N. Roscoe (Ed.), Aviation Psychology. Ames, Iowa: Iowa State University Press, 1980, 39-47.

Roscoe, S.N., Hull, J.C., Simon, P.M., & Corl, L. Human factors affecting pilot performance in vertical and translational instrument flight: Phase I interim scientific report. Las Cruces, NM: New Mexico State University, Department of Psychology, Behavioral Engineering Laboratory, Technical Report BEL-81-1/ONR-81-1, December 1981.

Vertut, J. Man-machine aspects in the control of the ASV. Commissariat a l'Energie Atomique, Institut de Protection et de Surete Nucleaire, Department de Protection Technique, Annual Report on contract period July 1982 - July 1983, June 27, 1983.

Whitney, D.E. Resolved rate control of manipulators and human protheses. IEEE Transactions of Man-machine Systems, 1969, MMS-10, 47-52.

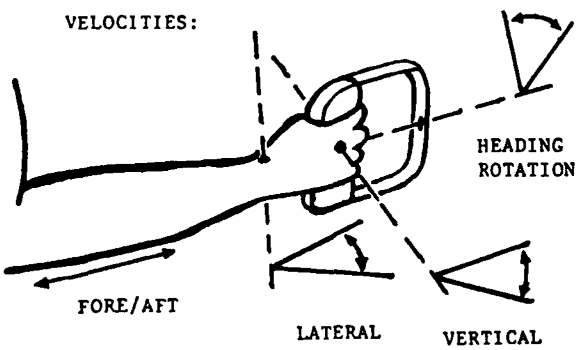


Figure 1. Rate controller configuration, 4 df.

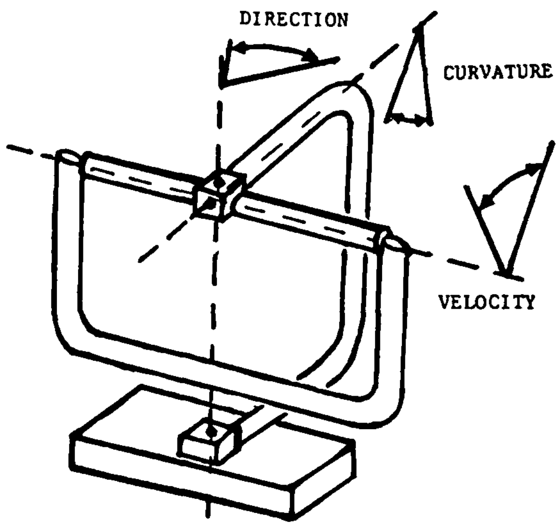


Figure 2. Vertut's 3-df controller configuration.

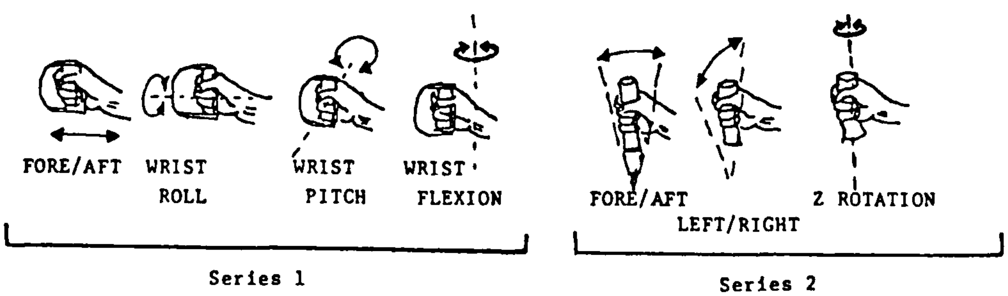


Figure 3. Dimensions of hand/arm movement.

Validation of the General Classification Test For Selection Of Commissioning From the Ranks Candidates

Captain Deborah A. Saudino
Canadian Forces Personnel Applied Research Unit
Willowdale, Ontario, Canada

Abstract

The General Classification test is used in screening all officer candidates for the Canadian Forces. This paper summarizes the research conducted evaluating the test's predictive validity in selecting candidates nominated under the Commissioning From the Ranks program. On-job performance, as measured by the Performance Evaluation Report, was used as the criterion. Results indicated that the GC correlated significantly with the intellect component of officer performance, as identified by a factor analysis, but not with other performance factors. Similar results were found using a comparison sample of officers-in-general.

Introduction

In assessing individuals for commissioning as officers, the Canadian Forces (CF) screens candidates using the General Classification (GC) test, a measure of general learning ability. The present study of the GC test in officer selection was conducted in response to questions and observations regarding the use of the test in the selection of candidates nominated for Commissioning From the Ranks (CFR). This highly competitive program is open to superior servicemembers in the rank of Sergeant and above who have 10 years or more of service, meet education requirements, and achieve a satisfactory GC score. Candidates nominated for the program have proven themselves as superior senior NCOs and are considered by their Commanding Officers to possess necessary officer-like qualities. This study attempted to establish the validity of the GC for predicting subsequent CFR performance and to compare obtained results with the validity of the test in predicting performance for a sample of officers-in-general.

At the time of its development by the Royal Canadian Air Force during WWII, the GC was found to be an effective predictor of success in both flying and technical training. Later studies have shown it to correlate with: CF tests of verbal, mathematical, clerical, mechanical and electronic aptitude which are used for other ranks trade assignment (Rampton, Skinner, & Keates, 1972); aircrew selection tests (Saudino, 1982); and a measure of aptitude for second language training (Amyot, 1983). A recent CF study (LeGras & Staples, 1983) reports a correlation of 0.75 between GC scores and total scores on the Weschler Adult Intelligence Scale - Revised (WAIS-R).

In addition to its correlation with other aptitude measures, the GC been found to predict training performance for both officers and other ranks candidates. Studies have demonstrated that the GC is positively related to success in other rank recruit (Mullin, 1978) and trades training (Ellis & Saudino, 1981), officer cadet university performance (Vandyke, 1982), and pilot training (Vandyke, 1982). It has also been shown to correlate with job performance ratings for other rank members (Simpson, 1982).

This wide range of information regarding the test's relationships with other aptitude tests and with training and job performance measures, for both

other ranks and officers, lends credibility to the acceptance of the GC test as both a measure of general learning ability or intelligence and as a useful selection instrument. Its use in officer selection is based on the assumption that a minimum standard of intelligence is required of all officers and that the more intellectually capable candidate has a greater likelihood of succeeding in officer training and in subsequently demonstrating superior on-job performance. Consequently, the same GC cut-off is applied for all commissioning programs, regardless of whether the candidate is being considered for acceptance through a civilian entry plan or through an in-service program. In recent years, little research has been directed at assessing the effectiveness of the test in predicting officer performance. The purpose of this study was to address this research issue.

Method

Sample. The CFR sample consisted of 250 Anglophone junior officers (Lts and Capts) selected for Commissioning From the Ranks in the years 1975 to 1982. A sample of 243 Lts and Capts was chosen at random from the officer population to serve as a "officers-in-general" comparison group.

Test. The GC is a 30-minute, 80-item multiple choice test used in the initial screening of both officer and other rank CF applicants. Separate norms and cut-offs are applied in the assessment of these two groups. English and French versions of the test are used as appropriate. The research summarized in this paper is based on the results from Anglophone samples only.

Criterion. A key consideration in this study was the choice of performance criterion to be used. In the validation of most selection tests, because of the ease of collection and standardization of scoring, training course results are commonly adopted as the performance to be predicted. The relationship of the GC to training results has been summarized above. For the CFR population, use of training information was considered inappropriate because of the lack of uniformity in the training of CFRs, due to differences in precommissioning ranks and officer classifications course structure. Furthermore, due to their years of experience in a related trade, the CFR pass rate on course is typically very high. This limited variability in course performance, particularly when only "pass/fail" results are available, would not yield the degree of discrimination between individuals necessary for test validation.

A more meaningful source of criterion data was deemed to be officer performance as measured by the annual Performance Evaluation Report (PER). The PER is a job performance measure that provides a readily available, standardized assessment for all officers, regardless of rank, classification, or specific employment. It rates the individual on 20 Performance Factors (PFs) and Professional Attributes (PAs), as well as on Potential for the next higher rank. A full list of the assessment categories is shown in Table 1. PFs and Potential are rated on a seven-point Likert-type scale, and PAs on a five-point scale.

Computerized annual PER data exist for all officers for the four year period, 1979-1982. This provided a total of 84 possible variables which could be used in the analysis, i.e., 21 PER assessment items X 4 years. Given the unworkable number of correlation coefficients this would produce (several of which would be expected to be significant by chance alone) the performance variables had to be combined or reduced. This was done by two means.

Table 1
CF Officer Performance Evaluation Report Items

Performance Factors		Professional Attributes		Potential
PF1	Acceptance of Responsibility	PA1	Professional Knowledge	Potential
PF2	Application of Knowledge	PA2	Appearance	
PF3	Problem Analyis	PA3	Physical Fitness	
PF4	Decision-Making	PA4	Conduct	
PF5	Preparation and Planning	PA5	Intellect	
PF6	Delegation	PA6	Integrity	
PF7	Oral Expression	PA7	Loyalty	
PF8	Written Expression	PA8	Dedication	
PF9	Performance Under Stress	PA9	Courage	
PF10	Cooperation			
PF11	Development of Subordinates			

The first means of combining the data was to average the ratings of the individual performance items over the four assessment years. This served to reduce the number of criterion variables to 21 and also increased the reliability of the ratings by minimizing the effects of error variance due to differences in raters, employment factors, experience in the job, etc. The second step was to factor analyse PER scores in order to identify the underlying components of officer performance being evaluated by this instrument. A common factor analysis using varimax rotation was conducted on the averaged results of the 1979 to 1982 PERs for a sample of junior officers (N=1802). This yielded six factors which accounted for 69.9% of the common variance. Table 2 identifies the factor names, together with the PER items loading on each, based on a factor weight inclusion criterion of 0.30.

Table 2
PER Items Loading on Each PER Factor

1. Job Performance:
   - Acceptance of Responsibity
   - Application of Knowledge
   - Problem Analysis
   - Decision-Making
   - Preparation and Planning
   - Performance Under Stress
   - Cooperation
   - Delegation
   - Professional Knowledge
   - Dedication
   - Potential

2. Intellect:
   - Oral Expression
   - Written Expression
   - Professional Knowledge
   - Intellect
   - Application of Knowledge
   - Problem Analysis
   - Preparation and Planning
   - Potential

3. Professionalism:
   - Conduct
   - Integrity
   - Loyalty
   - Dedication
   - Acceptance of Responsibility
   - Cooperation

4. Management of Subordinates:
   - Delegation
   - Development of Subordinates

5. Fitness and Appearance:
   - Appearance
   - Physical Fitness

6. Courage:
   - Courage

A rational examination of the factors shown in Table 2 suggested that the Intellect factor should be the component of officer performance which the GC would be expected to predict most strongly, followed to a lesser degree by the Job Performance factor. On the other hand, there was no theoretical basis from which to conclude that that the GC should have any relationship with the remaining factors (e.g., Courage, Fitness and Appearance).

Procedure. All available PER results for each CFR officer were averaged, yielding PER item scores which were based on rated performance over one, two, three, or four years, depending on time since commissioning. Factor scores were calculated for each individual, based on the averaged PER item scores and using the factor weights generated from the factor analysis conducted on the general junior officer sample. The six resulting scores were correlated with precommissioning GC scores yielding Pearson product-moment correlation coefficients. A similar analysis was carried out for the officers-in-general sample.

Results and Discussion

The correlation coeffients between GC and the six PER factors derived for the CFR and officers-in-general samples are shown in Table 3.

Table 3
Correlations Between GC Score and PER Factor Scores
For CFR and Officers-in-General Samples

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
CFR Sample (N=250)	r=.0469 ns	r=.1770 $p < .01$	r=-.0914 ns	r=.0181 ns	r=.0577 ns	r=-.0224 ns
Officers-in-General Sample (N=243)	r=-.0529 ns	r=.1661 $p < .01$	r=.0124 ns	r=-.0400 ns	r=.0792 ns	r=.0378 ns

As the above results indicate, the GC test was found to be significantly correlated with the Intellect factor for the CFR group as anticipated. The correlations for all other factors were not significantly different from zero. A similiar pattern of correlations was found for the officers-in-general sample, suggesting that the same relationship exists between GC test scores and performance ratings for officers selected from either population. Clearly, the GC is equally effective in the prediction of performance for both groups, negating the possible argument that the test may be of limited use in the selection of CFRs.

While the observed correlation between GC and the Intellect performance is significant, it is quite small. This is not unexpected however, in view of several possible confounding variables. First, the GC score distribution for the CFRs was restricted by the prior application of a selection cut-off. Correcting for this restriction of range, using test distribution information for all other rank members in the rank of Sgt or above, increases the correlation coefficient to 0.21. A similar correction for the officers-in-general sample yields a correlation coefficient of 0.17.

Additionally, the magnitude of the correlation coefficients obtained in this study was limited by the instruments themselves. The GC has been shown to have good reliability as determined by the split-half and KR-20 internal consistancy methods, yielding values of .88 and .90 respectively (Rampton, Skinner, & Keates, 1972). No reliability values for the PER have been established but it is known that measures of this sort, which are based on subjective assessments of a large number of raters over an extended period of time, are typically lower by comparison (Borman, 1978). These factors would establish upper limits to the possible correlation coefficients that might be obtained. Notwithstanding the noted methodological constraints, obtained results indicate that it is not unreasonable to employ the GC test as one of several assessment instruments for selecting officers for the CF Commissioning From the Ranks program.

References

Amyot, K.A. Continuous French Course (CFC) Serial 8202: Descriptive Data (Technical Note 14/83). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1983.

Borman, W.C. Exploring Upper Limits of Reliability and Validity in Job Performance Ratings. Journal of Applied Psychology, 1978, 63, 134-144.

Ellis, R.T. & Saudino, D.A. Selection and Trade Assignment (Men) Project (Working Paper 80-13). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1980.

LeGras, L.P., & Staples, R.E. A Correlational Analysis of the Relationship of the Weschler Adult Intelligence Scale - Revised with the General Classification Battery - English and the Canadian Forces Classification Battery (Personnel Selection Services Report 1983-1). Borden, Ontario: 1983.

Mullin, E.E.D. Project Profile: An Analysis of Male Recruit Attrition (CFPARU Report 77-6). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1977.

Rampton, G.M., Skinner, H.A., & Keates, W.A. Selection and Trade Assignment (Men) Project Status Report (CFPARU Report 72-7). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1972.

Saudino, D.A. Ajustment of GC Cut-off for ASC Candidates (Technical Note 2/82). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1982.

Simpson, S.P. The Performance Evaluation of Women in the Canadian Forces in Relationship to Supervisors' Attitudes Towards Women and the Theory of Self-fulfilling Prophecy: (Doctoral Disertation, Carleton University, 1982).

Vandyke, G.A. The General Classification Test (GC) as a Predictor of Success at Canadian Military Colleges and Basic Flying Training (Unpublished CFPARU Working Paper). Willowdale, Ontario: Canadian Forces Personnel Applied Research Unit, 1982.

Validity Generalization of Navy Selector Composites

Paul Foley and Leonard Swanson[1]
Navy Personnel Research and Development Center

INTRODUCTION

BACKGROUND AND PROBLEM

In recent years, Schmidt and Hunter (1977) and their colleagues have presented strong evidence that most of the variability in validity coefficients among various studies may be artifactual in nature. That is, the variability in observed validity coefficients across studies may be due to a number of factors, chief of which are sampling error, criterion unreliability, test unreliability, and restriction in range of ability of sample members. The validities of interest for validity generalization therefore are estimates of true validities. The mean and standard deviation of the true validities are obtained by correcting the mean and standard deviation of the observed validity distribution for criterion unreliability and restriction in range. The resulting distribution of estimated true validities is considered the Bayesian prior distribution whose mean, standard deviation, and credibility value (CV) represent the degree of validity generalization represented by that particular distribution of test composite scores. Rejection of the hypothesis of validity generalization for a distribution of validities requires that the 90% CV include zero.

APPROACH

Comparison of Validity Generalization Outcomes

The first analysis was based on 16 schools that had validity data on six or more samples for the General Technical, Clerical, Mechanical and Electronics classification composites of the Navy's Basic Test Battery which was replaced by the Armed Services Vocational Battery (ASVAB) in 1976. Two validity generalization analyses were conducted, one appropriate for validities uncorrected for restriction in range and the other appropriate for validities corrected for restriction in range. The purpose was to compare results obtained using a value representing an assumed average restriction in range (for uncorrected validities) and results based on individually corrected validities.

Validity Generalization and Occupational Groupings

The second objective of this effort was to investigate the effect of alternative groupings of Navy ratings on validity generalization outcomes. Four different grouping strategies were used based upon pooling validities across: (1) all Navy ratings, (2) the nine occupational groups in the Navy Occupational

[1]The authors express their gratitude to Dr. Kenneth Pearlman of the Office of Personnel Management for providing the computer program and to Mr. Greg Candell for conducting the various analyses.

Handbook (NOH), (3) the twenty occupational categories of the Navy's Official Classification Manual (CLASSMAN) and, (4) four BTB selector composites (BTBFAM) used for "A" school assignment.

RESULTS

Comparison of validity Generalization outcomes

For class "A" schools in each of the 16 Navy ratings, Table 1 presents the mean and standard deviation of the distribution of observed validities and the estimated mean and standard deviation of the distribution of true validities based on assumed average range restriction and on individually corrected validities. Also shown are the number of samples per rating and the 90% credibility values for both validity generalization procedures.

PLACE TABLE 1 HERE

The results in Table 1 support the concept of validity generalization across validity studies within each of the sixteen ratings (none of the 90% CV's include zero). For 15 of the 16 ratings, both validity generalization procedures yielded mean true validities of .43 or greater. All 15 also had 90% credibility values of .24 or greater indicating that in the future it is probable that validity results would be substantially above .24 in most schools. Validity generalization results based upon an assumed average range restriction were basically the same as those based on validities that had been individually corrected for restriction in range.

Validity Generalization and Occupational Groupings

Table 2 contains the average within-group estimated true standard deviation for distributions of validities (based on range corrected validities) for each combination of BTB composite and job grouping strategy and for the validities pooled across all Navy ratings. The last column presents the means of these values across BTB composites and the last row contains the means for rows 1-3. The values in parentheses are the average within group estimated true standard deviations for ratings randomly assigned to categories.

PLACE TABLE 2 HERE

The first hypothesis investigated was that the average estimated true standard deviation of selector composites observed within the three homogeneous job grouping strategies (NOH, CLASSMAN, and BTBFAM) would not be substantially less than the average estimated true standard deviation observed for the heteroge-

neous job category consisting of - all ratings combined. As can be seen, the average estimated true standard deviations for each of the four BTB composites were smaller for the three homogeneous job grouping strategies as compared to the heterogeneous - all ratings combined category. The one exception was for the GT composite which was slightly larger for the BTBFAM. The differences overall, however, were rather small. The hypothesis of no substantial difference between the average estimated true standard deviations was accepted indicating that differences between occupational groupings of ratings have little effect on the variation in validity coefficients.

The second hypothesis investigated was that the average estimated true standard deviation should be smaller for homogeneous occupational groupings as compared to the average estimated true standard deviation developed from the random assignment of ratings into the same number of categories as found within the NOH, CLASSMAN, and BTBFAM grouping strategies. As can be seen in Table 2 the average standard deviations for different job grouping strategies was typically smaller than those resulting from random assignment. This was true in eight of the twelve (three occupations groupings by four BTB composites) comparisons. The exceptions occurred with the GT and CLER composites for the NOH and BTBFAM groupings. Though moderating effects were found for groupings of ratings in the MECH and ELEC composites the differences between homogeneous and random groupings were not large.

CONCLUSIONS

Two validity generalization procedures were evaluated using validation data for BTB composites. One was based upon assumptions about the sources of artifactual variance (assumed average restriction in range) and the other based upon validities corrected for range restriction. Both approaches yielded similar results demonstrating the robustness of the validity generalization approach. Validity generalization results were obtained from five different rating grouping strategies which included a random category. It was hypothesized the variability of distributions of validity coefficients would be smaller for three homogeneous rating grouping strategies as compared with two heterogeneous groupings of all ratings combined and one comprised of ratings assigned randomly to categories. The differences between homogeneous and heterogeneous groups overall were slight suggesting that there is relatively little difference in the usefulness of the grouping strategies in terms of producing a significant moderating effect.

REFERENCES

Schmidt, F.L., & Hunter, J.E. Development of a General Solution to the Problem of Validity Generalization. Journal of Applied Psychology, 1977, 62, 529-540.

Table 1

Means and Standard Deviations of Observed Validity Distributions and Estimated Means, Standard Deviations, and Credibility Values of Estimated True Validities in 16 Ratings

Rating	N Samples	N Students	Observed Validity Mean	Observed Validity SD	Estimated true validities based on assumed average range restriction Est. Mean	Est. SD	90% CV*	Estimated true Validities based on range corrected validities Est. Mean	Est. SD	90% CV*
QM	6	2,783	.41	.073	.50	.055	.41	.55	.003	.55
SM	8	2,885	.24	.072	.33	.070	.24	.33	.096	.21
FT	7	2,952	.56	.072	.87	.101	.74	.78	.048	.72
TM	6	1,357	.29	.123	.47	.176	.25	.46	.172	.24
YN	12	5,036	.34	.090	.47	.109	.33	.49	.085	.38
SK	10	8,153	.44	.061	.58	.071	.49	.57	.043	.51
DK	10	1,624	.32	.078	.43	.047	.37	.43	.093	.31
EM	6	8,469	.42	.049	.61	.062	.53	.61	.065	.53
IC	6	4,038	.44	.076	.63	.098	.51	.60	.059	.53
HM	12	34,471	.58	.039	.73	.046	.67	.72	.035	.67
ET	9	9,698	.49	.053	.82	.083	.72	.72	.042	.67
ETR	7	4,865	.42	.119	.74	.202	.48	.60	.103	.47
ETN	8	6,079	.43	.091	.75	.150	.55	.62	.079	.52
OS	12	9,099	.41	.091	.65	.136	.47	.64	.064	.56
TM	8	1,463	.38	.049	.56	.000	.56	.55	.086	.44
PH	6	2,418	.49	.050	.64	.042	.59	.64	.000	.64
Mean							.494			.497

* 90% CV = Credibility Value

Table 2

Average Within Group Estimated Standard Deviations of True Validities Corrected for Restriction in Range for each BTB Composite by Job Grouping Strategy and by Random Grouping

Basis for Occupational (Rating) Grouping	BTB Composites GT	MECH	ELEC	CLER	Row Means
NOH	.093 (.091)	.087 (.112)	.111 (.115)	.080 (.079)	.093 (.099)
CLASSMAN	.073 (.083)	.069 (.090)	.062 (.091)	.065 (.071)	.067 (.084)
BTBFAM	.103 (.099)	.100 (.120)	.106 (.119)	.094 (.089)	.101 (.107)
All Ratings	.099	.119	.126	.094	.110
Column Means (Rows 1-3)	.090 (.091)	.085 (.107)	.093 (.108)	.080 (.080)	.087 (.097)

AD-P003 276

Improving Navigator Selection

Garry L. McDaniel
Academic Program Manager
ATC Instructor Training School

Abstract

The purpose of this paper is to present the results of the first phase of a three-phase investigation of U.S. Air Force navigators. The results of this study will be used by test developers, instructors, and administrators to improve selection, classification, training and utilization of navigators in the Air Force. The current study is a longitudinal research effort designed to track a single group of subjects through undergraduate, graduate, transition training and for one year in an operational assignment.

Background

The use of mental testing in selection and classification of candidates for navigator training has been in effect since 1941. Many research efforts designed to improve the effectiveness of the selection and classification process have been made. These include the development and revision of the Aircrew Classification Battery (ACB) and the Air Force Officer Qualifying Test (AFOQT). Special studies have been performed to identify the future role and training requirements for navigators through 1990 and to determine the causes of attrition in Undergraduate Navigator Training (UNT). There have also been several attempts by instructional and administrative personnel at UNT to identify and minimize factors contributing to the attrition problem. These efforts include curriculum and instructor evaluation, class sectioning, remedial math programs, and the use of a diagnostic pre-navigation screening test (Pre-Nav test).

The Pre-Nav test measures the level of navigator-related skills possessed by a student and includes test items on table reading, chart reading, instrument and scale reading, and basic math skills specifically relating to navigation. The Pre-Nav test differs from the AFOQT primarily on the basis of terminology used. For example, the quantitative items on the AFOQT are phrased in general terms while the similar items on the Pre-Nav test are specifically couched in the terminology of a navigation-like scenario.

The major advantage of the Pre-Nav test is that it provides a standardized measure for all students, since AFOQT scores are not available on everyone. The Air Force Academy does not use the AFOQT for selection or classification purposes. Civilians from the National Oceanographic and Atmospheric Association (NOAA) and students from the U.S. Navy and U.S. Marine Corps do not take the AFOQT. Also, scores may not be in the personnel records of some Air Force students, even though they may have taken the test.

Initial research with the Pre-Nav test indicated that an individual's Pre-Nav score could be strongly correlated to his or her chances for graduation or elimination. This correlation was determined by comparing UNT graduation/elimination against a Pre-Nav raw score frequency distribution. The Air Force Human Resources Laboratory, Pesonnel Research Division (AFHRL/MOAO) was able to validate these findings and research efforts were coordinated between AFHRL/MOAO and Mather AFB to improve navigator selection procedures.

One of the first steps towards obtaining a larger and more detailed data base which could be used to improve navigator selection was the development of the Basic Navigator Battery (BNB). The BNB was administered to all entering UNT students and consisted of the Pre-Nav test, and subtests consisting of Figure Analogies, Information Processing, Obstacles and Remedies and a Simulated Navigation Mission.

Because most of the previous studies were cross-sectional and limited in scope, it was deemed appropriate by Air Training Command to initiate a comprehensive, longitudinal research effort. This study differs from previous investigations in two major respects. First, it is a three-phase longitudinal study. Phase I, which this paper addresses, investigated the validity of the predictor variables using UNT performance measures as the criteria. Phase II will use performance measures from the advanced navigation courses (i.e., Navigator-Bombardier training, electronic warfare, etc.) and transition training. Phase III will investigate the correlation of the predictor variables with performance of the subjects at the end of their first year in an operational assignment. Phase III will be used to validate the navigator selection process against actual job performance.

Second, this study differs considerably in the choice of criteria to be used. All previous studies reported biserial correlations with a graduation/elimination criterion. In this study, a pass/fail criterion is used, but only for purposes of comparison. The major thrust of the Phase I research effort was to investigate all available measures of performance, to include academic scores, simulator results, and the results of actual flying missions. The purpose of this paper is to provide the findings and conclusions from Phase I of the three phase longitudinal study.

Approach

Only U.S. Air Force officers entering course N-V6A-F, Undergraduate Navigator Training (UNT), were considered as subjects for this study. The sample was further limited to UNT classes 80-18 through 81-17, which comprises one full year of Air Force officer input. UNT class 80-18 began training during January 1980. Each subsequent class began at two-to-three week intervals.

Subjects for the study were 591 Air Force officers entering Undergraduate Navigator Training during calendar year 1980. Because the sample will be used throughout the three phases of the longitudinal study and then used as a benchmark group for measuring improvements in selection and training, it is referred to as the UNT Reference Group. The demographic characteristics of this group are provided in Table 1.

Table 1. Distribution of Demographic Variables On the UNT Reference Group

Variable	Category	Percent
Sex	Male	96.2
	Female	3.8
Race	Insufficient data	
Age	Less than 23	14.9
	23	22.2
	24	18.3
	25	15.5
	26	10.8
	27	11.8
	28 or more	6.5
Rank	2nd Lieutenant	94.4
	1st Lieutenant	3.2
	Captain	2.4
Source	Officer Training School (OTS)	44.7
	AFROTC	29.1
	U.S. Air Force Academy	6.6
	UPT Eliminees (UPTE)	6.9
	Non-rated (NR)	3.7
	Non-Active Duty (NAD)	9.0

The primary purpose of this sample is to provide predictor and performance data during the three phases of the longitudinal study. It will also serve a secondary purpose as a control group for measuring the effects of any significant changes in selection or training procedures.

To determine the most effective methods of navigator selection, a variety of data was obtained on each student. This data consisted of ten predictor variables; five composite scores from the AFOQT, and the five variables from BNB. Performance variables from 15 intermediate UNT lesson scores and a UNT graduation/elimination variable were also obtained. A Pearson product-moment correlation matrix using the Bivariate Subsample Method was computed on each of the 10 predictor variables versus each of the 16 performance variables.

The AFOQT is administered to most UNT entries prior to selection and classification for training. The BNB was administered to the subjects during their first week of UNT training. The AFOQT scores, BNB scores and 16 UNT performance measures were matched on the subjects social security number.

Conclusions

Since the Navigator Selection Research Study is designed to be a three-phase longitudinal investigation, this paper does not provide major conclusions or recommendations. However, several observations regarding the findings in Phase I deserve mention. In general,

(1) the Navigator-Technical and Quantitative composites of the AFOQT are most valid of the five operational predictor variables,

(2) the Pre-Nav and Simulated Navigation Mission subtests are the most valid of the five experimental predictor variables,

(3) the Pre-Nav test is the most valid of the ten predictor variables used in this study, and

(4) correlation coefficients were highest with the didactic or classroom criteria, followed by the T45 simulator criteria, and then the flight criteria.

An additional finding resulting from Phase I is that changes to Form N of the AFOQT increased the predictive efficiency of the Navigator-Technical composite. However, even with the recent changes, the validity of the Navigator-Technical composite for predicting success in Undergraduate Navigator Training is not as good as it should be. Subsequent versions of the AFOQT test battery will incorporate changes based on the outcome of all three phases of this study.

An additional benefit from the research has been an increased concern towards improving recruiting strategies and preparing individuals interested in a navigator career prior to arriving at UNT. Several improvements have been made or suggested in these areas.

First, in the past the Nav-Tech portion of the AFOQT was administered only to recruits who expressed an interest in the navigation career field. This meant that of the approximately 40,000 personnel annually taking the test, only four or five thousand were taking the Nav-Tech AFOQT battery. Consequently, the navigator students gained through AFROTC and OTS were drawn from a pool of about ten percent of the potential. The AFOQT Nav-Tech battery is now given to all recruits. As these recruits begin to fill the student slots at UNT there should be an improvement in the quality as a result of having a larger pool from which to draw navigator candidates.

Secondly, the current recruiting strategy is one oriented toward selecting any candidate who meets a minimum standard (25th percentile on the Nav-Tech AFOQT battery). This strategy, in effect, assumes that a 25th percentile has as much potential as a very high percentile, which is of course not true. This recruiting strategy is easy to manage at the recruiter's level, however, for it means that they only have to find qualified (25th percentile or above) personnel who were interested in a navigation career. When previously the Nav-Tech battery was administered only to candidates who already had expressed an interest, half of the recruiter's search was done before the test. They could meet their quota with any qualifier.

AD-P003 277

Monetary Value of Navigator Selection Using the AFOQT

Bennie W. Roach
Air Force Human Resources Laboratory
Brooks AFB, Texas 78235

Abstract

The Navigator-Technical composite of the Air Force Officer Qualifying Test (AFOQT) has been used in the selection of candidates for Air Force navigator training for four decades. Expressing the value of the AFOQT in psychometric terms is often difficult for operational decision-makers to understand. A more appropriate way to express utility of the test is in terms of the "dollar criterion." This study uses the Schmidt-Hunter-Pearlman utility formula to compute the monetary value of two alternative selection strategies: (a) a zero utility approach which ignores test scores, and (b) the actual selection strategy which rank-ordered applicants based on test score. Results show that use of the Navigator-Technical composite percentile scores to select candidates for navigator training has considerable utility to the Air Force when expressed in monetary terms.

Introduction

Paper-and-pencil tests have been used by the Air Force to aid in the selecting of candidates for navigator training for more than 40 years. The Aircrew Classification Battery (ACB) was first developed for selecting aircrew members (pilots, navigators, and bombardiers) by the Aviation Psychology Program during World War II. During the early stages of the development of the ACB, selection for aircrew training was not based on predictor scores and most early classes had attrition rates of around 50%. Selection into Undergraduate Navigator Training (UNT) is now based on Navigator-Technical composite scores on the Air Force Officer Qualifying Test (AFOQT). Only applicants who have successfully passed a rigorous physical examination are considered. Additionally, all individuals selected for UNT must be college graduates. Since only highly restricted applicants are considered for selection and since current UNT attrition rates are only about 12%, the biserial correlation used to indicate the validity of the selection composite is quite low (.20).

From the standpoint of validity alone, it is difficult to defend the AFOQT as the primary instrument for navigator selection. Correcting the coefficient for range restriction, attenuation, and other statistical artifacts would increase the absolute value, but not the level of acceptance of the validity coefficient as an indication of the "value" of a selection instrument.

Dr. Ben Roach, AFHRL/MOAO, has proposed a strategy (Quota by Merit) where instead of looking for a minimum qualifier the recruiting system focuses on the high achievers of the entire pool and actively recruits them for a navigator career, even if they had showed no interest previously. He feels the recruiter who can identify an individual with a high potential for success as a navigator can use that very knowledge as a tool to sell the career to the candidate. In other words, if you tell someone he has a high potential for success as a navigator, the navigator career will seem more attractive to him. Concentrating on the high end of the spectrum gets away from the low potential qualifiers who happened to walk in the recruiter's door at the right time.

Third, at a recent ROTC Professor of Air Science (PAS) conference at Mather AFB, several suggestions were made to help ROTC students improve their skills prior to arriving at UNT. Individuals with technical backgrounds have tended to perform very well in navigator training, so an active, aggressive recruiting campaign for individuals with technical backgrounds was suggested. Also, at present, there is a ground school for pilots, but none for navigators. It was strongly recommended that a "rated" ground school be developed for ROTC students as the introductory information taught is beneficial to both pilots and navigators. The benefit would be to allow the "crew-concept" to be taught earlier, and fundamentals of basic flying such as aerodynamics, weather, plotting, use of the MB-4A computer, introductory navigation procedures, and ratio and proportion math problems relating to the flying environment (time, speed, distance and fuels) could be taught. Additionally, accurate navigation career information through films, briefings and motivational rides are needed to better inform ROTC and OTS students as to what a navigation career entails.

Efforts to improve navigator selection are continuing to be made. These efforts have focused on improving screening and selection of qualified personnel prior to arriving at Mather AFB for navigator training. As more information becomes available thorugh the Phase II and III research studies, selection and classification methods should result in a continued improvement to the AFOQT. Coupled with the other efforts discussed, the end result should be in students arriving at UNT with a better potential for a successful career in navigation.

An alternative to the use of the validity coefficient as a measure of the value of the AFOQT is to determine the utility of the test expressed in dollar terms. The use of monetary utility in personnel selection was first introduced by Brogden (1946) and improved upon by Cronbach and Gleser (1957). However, the Brogden-Cronbach-Gleser total utility formula had been largely ignored by personnel researchers for many years because of the complex equations and difficulty in determining SD_y, the standard deviation of job performance (Dunnette & Borman, 1979). However, a major breakthrough in the use of utility formulas occurred when Schmidt, Hunter, McKenzie, and Muldrow (1979) obtained rational estimates of SD_y. Using a global estimation procedure, Schmidt et al. used experienced supervisors as judges of subordinates' productivity. Their global estimation procedure suggests that SD_y represents a value that is about 40% of the annual wage of the employee.

The purpose of this report is to determine the utility of the Navigator-Technical composite as an instrument for selecting candidates for navigator training. Utility is expressed in dollar terms and considers both the cost savings due to decreased attrition in UNT and the dollar value expected from increased productivity of the new navigators. Monetary value is illustrated by using two alternative selection strategies: (a) a zero utility approach which ignores AFOQT test scores, and (b) the actual strategy used to select the cadets for training. The estimated monetary value of the actual selection strategy is determined by using the equation for comparing organizational interventions suggested by Schmidt, Hunter, and Pearlman (1982).

Approach

Subjects

The subjects for this study were obtained from an initial sample of candidates in UNT classes during FY80-82. Since the United States Air Force Academy discontinued use of the AFOQT in 1960, only UNT cadets commissioned through the Air Force Reserve Officer Training Corps and Officer Training School precommissioning programs were used. The final sample, which contained only cases with both valid AFOQT scores and UNT performance measures, was composed of 1,894 subjects.

Method

Determination of the monetary value of alternative navigator selection strategies using the AFOQT is a four-step procedure which includes:

(1) development of an equation which represents expected success rate,

(2) computer simulation of the probable outcomes of alternative selection strategies,

(3) determination of the terms used in the Schmidt-Hunter-Pearlman formula, and

(4) determination of the total monetary value of each strategy.

Step one, development of the expected success rate equation, is performed by first computing the actual success rate of the subjects at each scoring increment of the Navigator-Technical composite of the AFOQT. Since UNT cadets used in the sample had been selected, to a certain extent, based on performance on the Navigator-Technical composite, the number of cases at the low end of the scoring range was quite small. Therefore, success rate was determined for 10 scoring increments with each representing one decile of the total range of scores, i.e, 1-10, 11-20, 21-30, etc. A simple linear regression equation, using the midpoint of each scoring increment as the predictor and corresponding success rate for all UNT cadets who scored in that decile as the criterion, was used to determine expected success rates.

The second step was accomplished by using a previously developed computer algorithm (Roach, 1983). The specifics of the algorithm are too lengthy to be described in this paper, but basically the algorithm computes the probable values of several attributes for each strategy of interest (for example: the success rate of the candidates selected, mean percentile score on the predictor for the successful candidates, and required UNT input). Using z-transformation, the mean percentile group scores are converted to standard score form which can then be used in the Schmidt-Hunter-Pearlman formula. The required UNT input is a measure of the number of the selectees needed with each strategy in order to produce the desired number of graduates. Subtracting the desired number of graduates from the required UNT input yields the number of losses expected from the group of selectees. Multiplying the number of losses by the average training expense for each unsuccessful candidate yields the direct cost of attrition for the strategy. This is added to the cost of testing to provide the total cost of selection. The cost of selection is then used in the Schmidt-Hunter-Pearlman formula.

The third step requires determining the values for the terms used in the Schmidt-Hunter-Pearlman formula. This formula provides a way of assessing the economic impact of personnel programs on work force productivity and is defined as:

$$\Delta U_{1-2} = TN[(d_{t1} - d_{t2})\ SD_y - (C_1 - C_2)]$$

where ΔU_{1-2} = the difference in dollar value between two alternative strategies

T = the number of years duration of the training effect on performance

N = the number of desired graduates

d_t = difference in job performance in standard score units

SD_y = standard deviation of job performance in dollars

C = cost of selection

This formula is used to compute the total change in utility. The average utility of each graduate is computed by simply deleting the N in the formula. Standardized score, d_t, and cost of selection, C, were computed in Step 2, above. Tenure, T, is 5 since new navigators have a 5-year obligation after completion of UNT. The dollar criterion, SD_y, is \$8,211.65. SD_y is computed as .40 times the average annual wage of \$20,529.12.

The last step is determination of the total monetary value of each strategy. In this step, Strategy 1 uses test scores and Strategy 2 ignores scores. Therefore, the difference in dollar value between a strategy which uses test scores and the zero utility strategy is an estimation of the monetary value of the AFOQT using that particular strategy.

Results

Based on the success rates of the 1,894 subjects in the sample, the simple linear regression equation which represents expected success rate in UNT (Y') given the AFOQT score (X) is Y' = .7104 + .0028X. The frequencies and success rates at each decile and for the full range of scores on the Navigator-Technical composite is shown in Table 1.

Table 1

Distribution of Frequencies and Success Rates

Scoring Increment	Number of UNT Cadets: Pass	Fail	Total	Success Rates: Actual	Estimated
01 - 10	3	4	7	.429	.725
11 - 20	38	13	51	.745	.753
21 - 30	138	33	171	.807	.782
31 - 40	176	48	224	.786	.810
41 - 50	215	33	248	.867	.839
51 - 60	216	28	244	.885	.867
61 - 70	252	25	277	.910	.900
71 - 80	257	20	277	.928	.924
82 - 90	208	14	222	.937	.953
91 - 99	167	6	173	.965	.981
01 - 99	1670	164	1894	.882	.879

As Table 2 shows, the standardized score for the UNT graduates using the zero utility strategy was .096 and the cost of selection was estimated to be \$6,406.79 for each new navigator graduated. Since this strategy is used as the base rate against which the other strategy is compared, there is no monetary utility. The cost of selection using the actual selection strategy was estimated to be \$5,019.13, or a 22% reduction in the cost of selection using zero utility. The monetary value of this strategy using the Schmidt-Hunter-Pearlman formula was estimated to be 5[(.337 - .096) \$8,211.65 - (\$5,019.13 - \$6,406.79)] or \$16,833.33 per graduate.

Table 2

Summary of Monetary Value Analysis

Selection Strategy	Success Rate	Standardized Mean Score on Predictor	Cost of Selection Per Graduate	Average Utility Per Graduate
Zero Utility	.853	.096	$6,406.79	$ .00
Actual Selection	.882	.337	5,019.13	16,833.13

Discussion

This study uses the Schmidt-Hunter-Pearlman formula for comparing organizational interventions to determine the monetary value of the Navigator-Technical composite of the AFOQT for selecting UNT candidates. The estimated monetary utility to the Air Force can be substantial when AFOQT test scores are incorporated into the selection process. While the cost of selection is a one-time expense, the benefits obtained from selecting higher quality candidates is expected to occur throughout the duration of their tenure. The annual value of using the AFOQT for navigator selection during the most recent fiscal year is estimated to be $10 to $15 million over the 5-year period of obligation of 800-900 new navigators completing UNT each year. The major conclusion of this study is that the more efficiently AFOQT scores are used to select candidates for navigator training, the lower the costs of selection and the higher the utility to the Air Force expressed in monetary terms.

References

Brogden, H. E. On the interpretation of the correlation coefficient as a measure of predictive efficiency. Journal of Educational Psychology, 1946, 37, 65-76.

Cronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1957.

Dunnette, M. D., & Borman, W. C. Personnel selection and classification systems. Annual Review of Psychology, 1979, 30, 477-525.

Roach, B. W. Development of a decision support system for analyzing tradeoffs between alternative personnel selection strategies. Doctoral Dissertation, The University of Texas at Austin, August 1983. (To be published by Xerox University Microfilms.)

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. Impact of valid selection procedures on work force productivity. Journal of Applied Psychology, 1979, 64, 609-626.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. Assessing the economic impact of personnel programs on work force productivity. Personnel Psychology, 1982, 35, 333-345.

AD-P003 278

CURRENT RESEARCH IN ROYAL NAVY OFFICER SELECTION

by

ALAN JONES, Department of Senior Psychologist (Navy), MOD(UK)

The present Royal Navy Officer selection system derives from the World War 2 developments in the British Armed Forces. At that time serious manpower problems, particularly in the Army, led to the experimental and then executive use of extended assessment systems, which embodied many of the elements of the present day assessment centre (AC) approach. Although the most dramatic innovation was the introduction of group exercises and simulations, it should be emphasised that other applied psychological procedures were adopted, such as psychometric tests and more structured interviews.

The new style Royal Navy procedure was introduced in 1947, and, although there have been changes, the basic characteristics of the procedure are still recognisable. Candidates attend the centre for two days and are grouped for purposes of assessment into groups of five. Civilian candidates are aged between 17 and 25 years, and are in the top 20% of the population in terms of academic achievement. Between 1,500 and 2,000 candidates per year are assessed at the centre.

Their first day is spent completing a biographical questionnaire, taking 4 psychometric tests and a general and Service knowledge test, writing an essay, and being given a brief and practice on the types of group tasks they will face the next day. At the end of the day, the assessors individually review all the written evidence collected (except the psychometric test results), including reference reports from schools, universities and cadet forces. The assessment panel is usually composed of a President (Commodore or Captain, RN), a Naval Member (Commander, RN), a civilian Headteacher, and a Personnel Selection Officer (usually a Women's Royal Naval Service Officer). The composition varies slightly for Royal Marines and Women's Royal Naval Service candidates.

On the second day the assessors see the candidates for the first time. Candidates undertake two group exercises. One is a "command situation" where each candidate in turn is the nominated leader of the group, which has to meet certain objectives in a physical task. The other is a leaderless group discussion, where candidates are given a written scenario and a problem and told to come up with a group solution. Obviously both these exercises are intended to give the assessors an opportunity to see how each candidate interacts with the other members of the group: who appears most influential, who does not work with fellow members of the team to meet group objectives and so on.

After the group exercises there are two interviews, one with the Personnel Selection Officer and one with the other three panel members. The Personnel Selection Officer's interview concentrates on background factors (domestic circumstances etc), whilst the panel interview covers academic and spare-time activities, reasons for wishing to join the Service, and a number of specific areas (for example, specialisation preferences and encounters with the police).

Once the Personnel Selection Officer has reported back to the other panel members and the Headteacher has made some assessments on academic aspects, the collection of evidence is complete. Each assessor then individually reviews the evidence to arrive at an assessment of the candidate's suitability to enter

training. It must be emphasised that the centre's concern is with the identification of potential and primarily with the prediction of success in initial Officer training, completion of which is usually one to two years in the future. It is not concerned with the prediction of operational performance or promotion. When all assessors have completed this individual review, they each give a mark on a 0-9 scale which summarises their assessment. Candidates are then discussed in turn, assessors giving their own opinions and commenting on those of others, with the President always speaking last. Discussion continues until broad agreement is reached or it is recognised that differences in views are unlikely to be reconciled. Assessors then give a final mark which is aggregated to produce the centre's overall assessment of the candidate in terms of suitability to enter training.

Since the introduction of this form of selection a large number of studies of its validity have been carried out. Research carried out on entrants between 1968 and 1981 (total N = 2144) has shown average correlations of 0.29 between the centre's overall assessment (which is not divulged to trainers) and Professional examination results (in subjects such as Navigation, Engineering and Seamanship) at the Naval College, and 0.39 between the assessment and leadership ratings at the College. The Naval College examination results and the leadership ratings are combined to produce a final class of pass, with which the centre's assessment achieved an average correlation of 0.35. Examination of the individual studies shows an increase in the centre's predictive validity over the above time period. Research on entrants since 1981 has confirmed this trend (correlations of around 0.50 with class of pass at the College).

Table 1 shows the relationship between the centre's assessment and wastage during initial training (some 8 months in duration).

TABLE 1: CENTRE OVERALL ASSESSMENT AND WASTAGE IN ROYAL NAVY OFFICER INITIAL TRAINING (1974-81)

CENTRE ASSESSMENT	N	PERCENTAGE	
		COMPULSORY WASTAGE	VOLUNTARY WASTAGE
Doubtful Potential	258	18	13
Fair Potential with Some Shortcomings	336	12	12
Adequate Potential	216	10	11
Good or High Potential	243	6	14
TOTAL	1053	11	13

It can be seen from Table 1 that although the centre assessment achieved a moderate level of relationship with compulsory wastage (correlation around 0.35), there was no relationship with voluntary wastage. Research on individual parts of the procedure have also shown no relationship between these parts (tests, group exercises etc) and voluntary withdrawal, with the possible exception of the short test of Service knowledge. We have therefore embarked on a number of research projects to determine whether any instruments can be produced which can predict voluntary wastage.

Much of the voluntary wastage occurs very early on in training - within a month of entry to the College - and so one obvious approach is to consider this early withdrawal behaviour as the result of very dramatic mismatch between the entrants' vocational orientation and interests and Naval training as they experiences it. We were aware of the apparently successful use of the Strong Campbell Vocational Interest Blank in predicting wastage at the US Navy Academy, and also favourable results from work with the Canadian Forces. We therefore looked around for a suitable inventory for our use and came up with Holland's Vocational Preference Inventory; here examinees have to indicate their like or dislike for 160 jobs. We have carried out a number of studies on such aspects as candidates' "faking good", and achieved some prediction of voluntary and compulsory wastage using a profile scoring approach, as shown in Table 2.

TABLE 2: VOCATIONAL PREFERENCE INVENTORY PROFILE AND INITIAL TRAINING WASTAGE (1980-81)

PROFILE SCORE	N	PERCENTAGE	
		VOLUNTARY WASTAGE	COMPULSORY WASTAGE
LOW RISK			
3+	107	5	6
0-2	91	22	12
-1 to -2	53	17	13
-3 or less	65	28	20
HIGH RISK			
TOTAL	316	16	12

Table 2 suggests that the profile scoring was tapping some general motivational factor(s) since it was related to voluntary and compulsory wastage. Unfortunately an initial cross-validation did not show any relationship between profile score and wastage, but sample size was relatively small (N = 190, with only 16 and 24 cases of voluntary and compulsory withdrawal respectively).

Another approach has been the development of biographical predictor scores. The systematic use of biographical data has been largely ignored in the UK (in contrast to the USA). Using a set of 52 biographical items, we have constructed predictors of a number of aspects of initial training including voluntary wastage. At the moment cross-validation data are being collected, but in the derivation sample a weighted combination of 18 items correlated around .40 with voluntary withdrawal. Of course, some level of prediction would be expected in the derivation sample and the cross-validation will be the true test of the

score's value. Examination of the items included in the score suggests that entrants with more technical interests, an absence of failure in their academic history, some involvement in relevant cadet forces, and a reasonable level of spare-time activities are less likely to leave voluntarily.

This paper has restricted itself to a brief account of the current RN Officer selection procedure and of some of the research being undertaken to improve the prediction of voluntary withdrawal in training. There are, of course, many other areas of research activity, such as the analysis of how the panel of assessors reach their final assessment and developments in aircrew selection, but there is not time to describe them here. It is, however, hoped that the reader has obtained some idea of the procedures and research efforts taking place in the United Kingdom.

AD-P003 279

Air Force Academy Measures of Performance as Predictors of Promotion Potential

Colonel Kenneth H. Fleming
United States Air Force Academy

Abstract

A recurring question with respect to performance is the validity of the standards of measurement. The USAF Academy, along with many other institutions, uses various standards of performance to measure the achievement of cadets while they are at the Academy. Since almost all cadets are constrained to enter the Air Force upon graduation, there is a continuous chronological record of performance. Record keeping is extensive and identifiable at a personal level, so that the Academy and the follow-on U.S. Air Force environment are ideal institutions for the investigation of specific performance measures. In particular, is there any correlation (causality) between measures of performance prior to, during, and subsequent to the four years at the Academy. Knowledge of these relationships (or lack) would be valuable in validating old standards or developing new ones.

Introduction

This study focuses on the first twelve classes who have graduated from the Academy (1959-1970); these twelve classes have all had a recorded opportunity to compete for selection to major in the U.S. Air Force. Only those officers who remain in the service are considered for this study. The data sample contains 2072 officers who have competed for major and 975 who have competed for lieutenant colonel. These results apply to only Academy graduates as compared to other Academy graduates.

The Air Force promotion system is tied to years in service. Beginning with the rank of major, however, a "best" qualified standard applies and there is provision for early promotion. For a typical year group, five percent will be promoted early, eighty percent on time, and fifteen percent will fail at selection in the primary zone; of these, approximately five percent will be late-selected. The same setup is used for promotion to lieutenant colonel except that the percentage rates for achieving the promotion are smaller. Because of data limitations, this study makes no distinction between early promotions. That is, a three-year-early promotion is treated the same as a one-year-early promotion. An-on time promotion is self-explanatory, and failure to be promoted in the primary zone is considered a pass-over regardless of subsequent selection.

The Model

A straightforward achievement model is proposed. Earlier measures of performance are thought to "identify" unobserved characteristics which make success more likely. Success is defined as selection for below-the-zone or on-time promotion to major or lieutenant colonel. Early, on-time, or non-

selection to major and lieutenant colonel are the dependent or success measures. High school, Academy, and Air Force performance measures are the independent or causal variables. If our performance measures are valid then we should be able to discern statistically significant effects on the probability of promotion. All officers are constrained to fall into only one of the three categories so that the dependent variable is trichotomous and exclusive. OLS was used to estimate the trichotomous model and results indicated that so many officers were promoted on time that the independent variables added little explanatory power to this class. There were, however, large and significant effects on the probability of early promotion to both major and lieutenant colonel. The most interesting results were therefore estimable with a dichotomous dependent variable--the probability of below-the-zone promotion versus not being selected. A logistic model was used to estimate the relationship. Figure 1 gives a two-dimensional picture of this procedure.

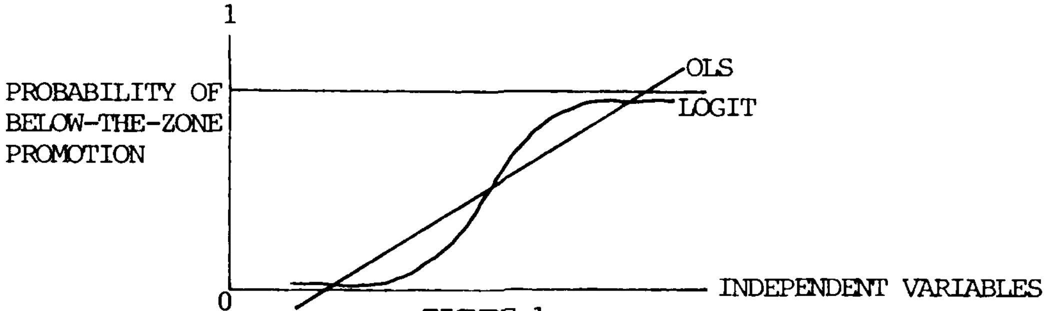


FIGURE 1

The basic features of the logistic model are interaction between the independent variables, largest effect at the midpoint, and an asymptotic approach to zero or one in the probability sense. Ordinary least squares was also used to estimate the parameters. OLS presents some difficulties as Figure 1 makes clear. The ordinary least squares procedure fits a straight line to the X observations. Thus, for extreme values of X_1, it is possible to achieve probabilities less than zero or greater than one. The additive form of OLS also does not allow for interaction between the independent variables. The logit model results are compared with OLS estimates.

It must be emphasized that this type of model is explanatory and not predictive. Thus, we cannot predict the performance of any one individual based on this model. Instead we are investigating the average effect of the independent variables; that is, how the mean value estimated for hundreds of individuals affects the mean probability of promotion.

Variables

The variables for each individual are as follows: parents' status (military or civilian), verbal aptitude and math aptitude scores upon entry to the Academy, order of merit academically and militarily at the Academy, rated status in the Air Force (pilot, navigator, or support), academic level of achievement in the Air Force (college, master's, or PhD), military schooling in residence (Squadron Officers School or Air Command and Staff) and whether or not the officer has combat experience. Verbal and math aptitude are

entered as the score achieved on the college entrance exam. Order of merit is entered as a class adjusted percent; for example, if an individual finished first in a class of 500 this would be recorded as 1 - 1/500 or .998. The other variables are entered as binary variables--one if the individual is in the class, and zero if he is not. The excluded classes are support, college degree only, and non-resident military education. Other candidate variables are insignificant and do not appreciably affect the value and statistical significance of the selected variables whether included or not.

Results

The largest and most statistically significant coefficients to major were the individual's Academy standing in order of merit leadership (OMLDR), order of merit academic (OMACA), and possession of an advanced degree; this result indicates that the two Academy performance standards are outstanding measures, at least on the average, of the ability to succeed in the Air Force. OMLDR is larger than OMACA and is therefore a better explanatory variable for below-the-zone promotion to major. However, it is important to keep in mind that both are positive, large, and highly significant. Parent's occupation was entered as either military or not--it was negative (-4% to major and -2% to lieutenant colonel) and highly significant. This means that for below-the-zone promotion it is helpful to have parents in the military--a result which is in accord with our intuition. Both the coefficients for a master's degree and for a PhD were positive and highly significant. Both were measured as the difference between these degrees and a bachelor's degree so that there is essentially no difference between the master's and the PhD. On average, the possession of any advanced degree will increase the chances of below-the-zone promotion, relative to a bachelor's degree, by about 7.5 percent in the OLS formulation and 8.5% in the logistic model (other variables held at their mean).

Verbal and math aptitude had conflicting signs, but neither was significant at a high level, so it is likely that their effects were subsumed by other variables. This hypothesis was tested by running the regression with and without the aptitude scores. The only appreciable change in any of the variables was in the constant term. For the logit model with the aptitude scores the constant term itself gave a probability of below-the-zone selection of 10%; without the scores the probability was 8.9%. Clearly the second value is more consistent with our a-priori beliefs. The difference appears small, but it must be remembered that the logit model becomes asymptotic at the extremes, so a 10% difference is appreciable. The same calculation was run with the OLS model with more pronounced results. The math and verbal aptitude scores were directly related to the constant term--no other explanatory variables were appreciably changed. For this reason the aptitude scores are excluded in further calculations.

Consider an extreme specific example for below-the-zone promotion to major with values taken from the logistic model. An individual who was first in his class in both OMLDR and OMACA, has parents in the military, is a pilot, has had combat experience, holds a master's degree, and has completed SOS by correspondence will have a 46.8% probability of making a below-the-zone

promotion--this is 42% above the Air Force average and 32% above the average in this data sample! An individual with the same qualifications who finishes last in his class in both OMLDR and OMACA, will have an 18% chance of a below-the-zone promotion--take away the master's degree and this drops to a 9% probability. The figures for the OLS estimate are equally impressive--37% versus 7% and zero without the master's degree.

It is clear that the overwhelmingly significant predictors are OMLDR, OMACA, and advanced academic degrees. For our extreme example each ten percent drop in OMLDR will produce approximately a 2.5% drop in probability of below-the-zone for the logistic model, other things held equal. Table 1 presents various representative combinations of characteristics for the logistic model. Table 2 presents the same data for the OLS model. It is difficult to choose between the models. The logistic model has a better theoretical basis, but it appears that it may be overestimating probabilities at the lower end. The models differ as to the precise values of the coefficients; however, they both are positive, in the same range, and highly significant.

Tables 3 and 4 give the results for lieutenant colonel. Again OMACA, OMLDR, and advanced degrees were strong positive predictors. However, OMLDR decreased in absolute value and OMACA increased. They are almost equal in value and the combined effect is greater than the combined effect to major. It appears that some individual outliers have been eliminated at selection to major and that the Academy performance measures continue to strongly influence promotion performance. Further evidence for this comes from the small increase in R^2--the model is doing a "slightly better" job of predicting individual behavior. The two estimating procedures show basically the same results with the logit model again giving higher estimates at all levels. It appears that the "interaction" effect of the logit model raises the probability estimate at all levels--this may actually be a better estimate at the higher probability levels since promotion boards consider the "whole person" when selecting for below-the-zone. The lieutenant colonel figures show a large and statistically significant negative effect for rated navigator status versus pilot or support. The model was run with master's degree as the dependent variable. There was a significant relationship between this and academic order of merit, but no relationship for leadership order of merit. This result confirmed an indirect effect for academic order of merit; this effect turned out to be small for major (1%), but larger for lieutenant colonel (3%).

Conclusions

The academic and leadership performance measures used at the Academy are excellent predictors, on average, of promotion potential. Leadership is larger for promotion to major, but both are equal to lieutenant colonel. An advanced degree is the other large predictor. Rated status is neutral to major, but negative and significant for navigators to lieutenant colonel. Combat and parental status are both significant but small contributors, and military school in residence is not significant.

TABLES

GIVEN: MASTER'S DEGREE; NO SOS IN RESIDENCE; NO COMBAT; NO MILITARY PARENTS

TABLE 1 (LOGIT RESULTS) | TABLE 2 (OLS RESULTS)

TABLE 1: PERCENT OF CLASS STANDING OMLDR	OMACA	PROBABILITY OF BELOW-THE-ZONE TO MAJOR: PILOT	NON-RATED*	TABLE 2: PERCENT OF CLASS STANDING OMLDR	OMACA	PROBABILITY OF BELOW-THE-ZONE TO MAJOR: PILOT	NON-RATED*
1	1	38.4	34.9	1	1	31.8	29.1
.75	.75	30.6	27.5	.75	.75	24.7	22.0
.5	.5	23.8	21.2	.5	.5	17.6	14.9
1	0	28.6	25.6	1	0	22.6	19.9
.75	.25	26.1	23.4	.75	.25	20.1	17.4
0	1	19.7	17.4	0	1	12.5	9.8
.25	.75	21.7	19.3	.25	.75	15.0	12.3
0	0	13.6	10.0	0	0	3.4	.7

*Note: There is no statistical difference between navigator and non-rated. The figures should be interpreted as the difference between the average of the sample (14.5%) and the given figures.

GIVEN: MASTER'S DEGREE; NO SOS IN RESIDENCE; NO COMBAT; NO MILITARY PARENTS

TABLE 3 (LOGIT RESULTS) | TABLE 4 (OLS RESULTS)

TABLE 3: PERCENT OF CLASS STANDING OMLDR	OMACA	PROBABILITY OF BELOW-THE-ZONE TO LT COL: PILOT	NAVIGATOR	TABLE 4: PERCENT OF CLASS STANDING OMLDR	OMACA	PROBABILITY OF BELOW-THE-ZONE TO LT COL: PILOT	NAVIGATOR
1	1	44.5	25.9	1	1	34.7	19.6
.75	.75	35.8	19.6	.75	.75	26.8	11.6
.50	.50	28.1	14.5	.50	.50	18.9	3.7
1	0	29.4	15.3	1	0	20.3	5.1
.75	.25	28.8	14.9	.75	.25	19.6	4.4
0	1	26.9	13.8	0	1	17.6	2.4
.25	.75	27.5	14.2	.25	.75	18.3	3.1
0	0	16.0	7.7	0	0	3.3	$-11.9 \cong 0$

Note: For promotion to lieutenant colonel navigator status is negative and highly significant. There is no statistical difference between pilot and non-rated. The figures should be interpreted as the difference between the average of the sample (17.0%) and the given figures.

AD-P003 280

THE EFFECTIVENESS OF BIOGRAPHICAL INVENTORIES IN PREDICTING SUCCESS IN OTS

Deborah L. Rogers
Manpower and Personnel Division
Brooks Air Force Base, Texas 78235

Abstract

Biographical inventories have been used in officer selection and classification tests since the beginning of World War II with the advent of the Aircrew Classification Battery. These inventories, used to predict success in officer, pilot, and navigator training, were used on successive forms of the Aircrew Classification Battery. Officer and pilot biographical inventories were also used on the Air Force Officer Qualifying Test (AFOQT) until Form O, the current operational form, was developed. However, in reviewing the history of the inventories, it appears that the inventories were not included in an analysis of the composites used to predict success in officer or pilot training. The current study was designed to see if the inventories improve the predictiveness of the Pilot and the Officer Quality (Academic Aptitude) composites of the AFOQT.

BACKGROUND

The use of biographical inventories to select and classify officers dates back to the Aircrew Classification Battery (ACB) developed at the beginning of WWII. The seventh version of the ACB, July 1943, contained the first biographical inventory subtest for pilots and navigators. These subtests, referred to as Biographical Data Blanks, were thought to have "appreciable validities and afforded a way of taking students' backgrounds into consideration in their assignment." (DuBois, 1947, p. 96).

The Biographical Data Blanks for pilots and navigators were used in the ACB until the end of the war when the need arose to revise the test for use on the post-war population (the Biographical Data Blanks on all wartime forms of the ACB were normed on WWII aircrew applicants, and correlations computed on these subtests were based on data gathered from pilot training only). The February 1947 ACB was implemented for this purpose.

The biographical subtests of the February 1947 ACB were not actually revised, but the "Biographical Data, Navigator" was renamed "Biographical Data, Bombardment-Observer." There was work done during this time to develop "...biographical inventories to predict success in the Air Force Officer Candidate School and success on the job during the first six months following graduation" (Dailey & Gragg, 1949, p. 54). All biographical subtests were normed on pilot trainees, and validated against elimination from pilot training. It was subsequently found that the validity of the Biographical Data Blanks "...varied greatly in validity against the criterion of graduation vs elimination for lack of motivation" (Dailey & Gragg, p. 53). For this reason the biographical subtests were revised in hopes that they would be more valid than the old ones.

The February 1947 ACB was revised in April 1951. At this time the Biographical Data Blanks for pilots and navigators were merged making a single biographic subtest for both pilots and navigators. These subtests were normed on aviation cadets and aviation applicants.

The Aviation Cadet-Officer Candidate - Qualifying Test (AC-OC-QT) AXA and AX[illegible], an experimental test, was introduced in 1951 as another attempt to better predict success in Officer Candidate School (OCS). In this test, the Biographical Data Blank was revised to include 50 percent pilot items and 50 percent items selected "on an *a priori* basis as predictive of officer success" (Tupes, 1953, p. 1). It was later found that the officer items had little validity when predicting OCS success and on-the-job performance. The Biographical Data Blank was not scored as part of the Officer Quality composite. The AC-OC-QT was normed on male members of OCS and was validated against graduation/elimination from OCS.

In 1951 the AC-OC-QT became part of a preliminary version of the Air Force Officer Qualifying Test (AFOQT) Form A; however, the officer biographical items were not carried forward. In Form A an entire booklet made up of 130 newly developed officer biographical items was devoted to a Biographical Information subtest. The Biographical Information subtest was used in the Officer Quality and Pilot composites. Form A was administered to OCS cadets as well as Air Force Reserve Officer Training Corps (AFROTC) cadets. Form B of the AFOQT contained a modified version of the Biographical Information booklet including 20 additional items. The implementation of this test marked the beginning of selection for the newly developed Air Force Academy (AFA), and was normed on all three commissioning sources OCS, AFROTC and the AFA.

Form C contained a separate biographical inventory for the Officer Quality composite and the Pilot composite. The Pilot Biographical Inventory (PBI) remained unchanged through AFOQT Form G. The number of items in the Officer Biographical Inventory (OBI) was almost tripled with the advent of Form D (150 items). The number of items of the OBI remained unchanged through Form G (Valentine & Creager, 1961).

AFOQT Form G was used until 1964 when it was revised and named AFOQT-64. The OBI and PBI were carried forward, but the number of items was reduced. The OBI was not administered to female applicants as the subtest was normed on male AFROTC cadets and was not thought to be suitable for female applicants. Neither the OBI or the PBI was included in reliability and validity analyses of the Officer Quality and Pilot composites (Miller & Valentine, 1964). These subtests remained virtually unchanged through AFOQT Form M. When AFOQT Form N was implemented the OBI was removed from the Verbal & Officer Quality composites. It was removed because:

> The OBI was previously taken only by males since it was composed of activities associated with males. Unable to remove sex bias from the items, the subtest was eliminated from the revised AFOQT (Gould, 1978, p. 8).

The PBI was replaced with the Pilot Biographic and Aptitude Scale. The norming sample for Form N included cadets from Officer Training School (originally OCS), AFROTC, the AFA, basic airmen and a sampling of second lieutenants.

When the current operational form (Form O) of the AFOQT, was implemented, a working group decided to eliminate the Pilot Biographic Inventory and Attitude Scale because of low validities and possible racial and sex bias of the subtest.

There has been to date no conclusive validity study on biographical inventories encompassing variables such as race, sex, and success in OTS and pilot and navigator training. As mentioned, the most recent studies done on the AFOQT did not include the biographical subtests in their analyses. Therefore the purpose of this study is to measure the predictiveness of the OBI and PBI, and the Officer Quality (now referred to as Academic Aptitude) and Pilot composites with and without the corresponding biographical inventories.

METHOD

The selection of subjects for this study was limited by two constraints: (1) source of subjects and (2) availability of predictor data. There are currently three sources for commissioning in the U. S. Air Force; OTS, AFROTC, and the AFA. Completion of precommissioning training at AFROTC and the AFA is contingent upon successful completion of college training. In order to prevent confounding the findings, only selectees for training in OTS were used.

The second limitation was that the subjects must have AFOQT scores from Forms L, M or N on the Academic Aptitude composite and the Pilot composite and their respective biographic subtests available in the data files maintained by the Air Force Human Resources Laboratory. After imposing both constraints, the initial pool of subjects for this study was 3861 (105 Form L, 2031 Form M and 1725 Form N) individuals who were selected for training in OTS from FY75 through FY80. Among this sample there were 241 pilots.

The predictor variables used in this study come from the AFOQT Forms L, M and N. The Pilot composite and PBI (available in all three forms) were compared in one part of the study. The Academic Aptitude composite including the Officer Biographic Inventory were available only for Forms L and M, and are compared in another part of the study. The criterion variables used in this study were pass/fail and final course grade in OTS, and pass/fail in Undergraduate Pilot Training (UPT).

Pearson product-moment validity correlations were computed in the pilot subsample for the Pilot composite, the PBI, and a combination of PBI and the Pilot composite, separately for Forms L, M, and N. In the OTS subsample, validity coefficients were obtained for Academic Aptitude only, OBI only, and a combination of the two, separately for pass/fail and final course grade, within each of Forms L and M.

RESULTS

When analyses were done on the biographic inventories implemented during WWII the correlations were computed separately for the biographic inventory subtests versus success in flying training. The resulting correlations ranged from .01 to .29 and were not considered significant. When the AFOQT was implemented, the biographic subtests were not included in the analysis of the composites against success in flying training or officer school training. A purpose of this study was to validate more recent versions of biographical

inventories. Correlations were computed using the PBI only versus success in flying and precommissioning training. The results were similar to those obtained during WWII, with uncorrected correlations ranging from -.02 to .34. Correlations were also computed in the pilot subsample, using raw scores of the Pilot composite with and without the PBI (See Table 1).

Table 1

Correlation of Pilot Composite with and without the PBI versus Pass-Fail in UPT for Total Group

	Pass/fail Ratio	Pilot w/PBI	N	Pilot wo/PBI	N	PBI Only	N
AFOQT Form L	67/23	.32	12	.38	12	-.04	12
AFOQT Form M	83/17	.25	152	.28	152	-.02	152
AFOQT Form N	50/50	.41**	77	.36**	77	.34**	77

**p .01

Overall the correlations were not significant except for Form N where the PBI alone predicted success in UPT fairly well. The Pilot composite predicted success in UPT better without the PBI for both AFOQT Form L and M.

Table 2

Correlation of Academic Aptitude Composite with and without Officer Biographic Inventory Versus Pass-Fail in OTS

	Pass/fail Ratio	Pass-Fail in OTS[2]	N	OTS Final Course Grade	N
AFOQT Form L	93/07				
Acad. Apt. + OBI		-.02	105	.27**	97
Acad. Apt. - OBI		-.13	105	.16	79
OBI Only[1]		-.16	84	-.08	79
AFOQT Form M	88/12				
Acad. Apt. + OBI		.05*	2031	.48***	1789
Acad. Apt. - OBI		.01	2031	.46***	1409
OBI Only[1]		.07*	1582	.14**	1409

*p .05
**p .01
***p .001
[1]Females did not take the OBI, therefore the N's for OBI only are smaller than Acad. Apt + OBI and Acad. Apt - OBI.
[2].82 Pass; .18 Fail. Variables were coded: Pass = 1; Fail = 0.

Correlations were also computed for the OBI alone, and the Academic Aptitude composite with and without the OBI (See Table 2). The correlations for the OBI ranged from -.16 to .07 against pass-fail in OTS. Because the OBI was not included in previous validity studies, there are no comparative correlations. The OBI, when added to the Academic Aptitude composite in the OTS subsample, increased the prediction by only a small amount. On Form M the OBI alone was a fair predictor of OTS Final Course Grade. The items in the OBI, however, were made up of activities associated with males, and the bias could not be removed. This subtest was subsequently dropped from AFOQT Form N.

CONCLUSIONS

As a result of this study three conclusions can be drawn:

1. As with the biographic subtests used during WW II, the validity coefficients of AFOQT biographical inventories are statistically significant in some small subgroups, although these inventories do not appear to improve prediction enough to warrant their inclusion in a prediction composite.

2. The subtests were modified several times throughout the history of the development of the AFOQT in an effort to increase the validities and reduce possible bias. These efforts were unsuccessful.

3. The biographic subtests were not included in the AFOQT Form O because of the low validites and the inability to eliminate bias in the biographic subtests.

REFERENCES

Dailey, J. T., & Gragg, D. B. Postwar research on the classification of aircrew. Research Bulletin 49-2. Lackland AFB, TX: Directorate of Personnel Research, Human Resources Research Center, Air Training Command, November 1949.

Dubois, P. H. The classification program. AAF Aviation Psychology Program Research Report No. 2. Washington: U. S. Government Printing Office, 1947.

Gould, R. B. Air Force Officer Qualifying Test Form N: Development and standardization. AFHRL-TR-78-43, AD-A059 746. Brooks AFB, TX: Air Force Human Resources Laboratory, Personnel Research Division, August 1978.

Miller, R. E., & Valentine, L. D., Jr. Development and standardization of the Air Force Officer Qualifying Test - 1964. PRL-TDR-64-6, AD-600 782. Lackland AFB, TX: Personnel Research Laboratory, Aerospace Medical Division, March 1964.

Tupes, E. C. The validity of the Aviation Cadet - Officer Candidate - Qualifying Test AXA and AXB for prediction of success in USAF Officer Candidate School. Technical Report 53-35. Lackland AFB, TX: Personnel Research Laboratory, Human Resources Research Center, Air Research and Development Commmand, December 1953.

Valentine, L. D., Jr., & Creager, J. A. Officer selection and classification tests: Their development and use. ASD-TN-61-145, AD-269 827. Lackland AFB, TX: Personnel Laboratory, Aeronautical Systems Division, October 1961.

USAF-CANADIAN FORCES EXCHANGE OF OCCUPATIONAL ANALYSTS

LtCol J.L. Mitchell, USAF, & Capt Frank Strickland, CF
Occupational Analysis Program
USAF Occupational Measurement Center
Randolph AFB, Texas 78150

ABSTRACT

An exchange program was established in 1980 between the US Air Force and Canadian Forces occupational analysis programs. The initial two-year assignments were begun in the summer of 1981, with the expectation of just a direct exchange of working level occupational analysts. In the summer of 1983, the second round of the exchange was made. The program has been very successful; it has not only achieved its original objectives but there have been some very beneficial serindiptious results.

BACKGROUND

IN 1978, during a visit to San Antonio by members of the Canadian occupational analysis (OA) program, there was some discussion with the USAFOMC staff about an exchange program to provide a continuing dialogue between the USAF and Canadian OA programs. This was during the first Occupational Analyst's Workshop which brought together occupational anlaysts of the US Armed Services and those of Allied military services (Australia, Canada, and the Federal Republic of Germany) as well as representatives of other governmental agencies, civlian firms, and a number of universities. At the second annual Workshop at Randolph AFB in the spring of 1979, the CF-USAF exchange proposal was again discussed. In the following months, a specific plan was submitted through both the Canadian and United States governments for the exchange of one officer of each country to work in the occupational analysis program of the other nation.

The Canadian Forces occupational analysis program is under the Director Military Occupational Structures (DMOS) in the National Defence Headquarters at Ottawa, where it serves to survey jobs in all commands within the CF. The Canadian OA program is within the Defence Headquarters structure and its members study occupational data to make recommendations on trade standards, recruiting, training, and utilization policies. Since the Canadian Forces are unified, their OA projects involve surveys of land, sea, and air operational jobs. Occupational analysts work in teams with senior subject matter specialists to develop job inventories through field visits, administer the inventories at major bases, and analyze the results of the surveys. An Applications Officer on each team staffs and briefs the results to upper levels of management and to the operatioal commands, including training command, if appropriate.

The USAF OA program is one of the four major missions of the USAF Occupational Measurement Center, a major command special activity of the

Air Training Command (ATC). USAFOMC is functionally alligned within the Hq ATC Deputy Chief of Staff Technical Training operation. The USAF program normally involves only Air Force occupations. The professional staff specializes as inventory developers or occupational analysts, with some individuals who work officer studies and special management applications projects (including interservice surveys) fulfilling both roles. Studies may involve one, two or even several AF specialties, depending on whether the project objective is to evaluate a proposal to merge occupational fields, or is an in-depth look at specialty training requirements. Inventory Developers visit Air Force operational units to conduct small group interviews with senior technicians of a specialty. Analysts typically work fairly independently in analyzing the survey data and in developing an extensive Occupational Survey Report (OSR); they brief their own findings and recommendations up the chain of command to the Hq Air Training Command (ATC) Deputy Chief of Staff for Technical Training, the requesting command or agency, and to the appropriate Technical Training Center.

Both the Canadian and the USAF OA programs use the Comprehensive Occupational Data Analysis Programs (CODAP) developed by the Air Force Human Resources Laboratory (Carmark 1969; Morsh 1964; Christal 1974; Mitchell 1983). USAFOMC utilizes CODAP on the AFHRL UNIVAC system. This joint use of their computer makes possible the almost immediate implimentation of any new occupational research technologies developed by the Manpower and Personnel Research Division (MO) of AFHRL. The Canadian Forces OA program uses a CODAP adapted from the US Navy IBM version. In addition, the Canadian program has some special requirements, such as the use of bilingual inventories and somewhat different rating scales, which required further CODAP modifications. DMOS programmers have also developed excellent quality control and analysis programs to meet the unique CF needs.

Thus, while both systems operate with the same basic approach and computer programs, they are adapted to somewhat different operational situations. The main objective of the exchange proposal was to permit both to become fully familiar with the other's operations, and to bring back to their own program useful ideas, new techniques, and a thorough knowledge of newly developed computer programs.

THE INITIAL EXCHANGE

With the final agreement between the two governments in late 1980, the stage was set for the initial exchange of occupational analysts in the summer of 1981. Major (the Captain) Ian Falle, an aircraft maintenance engineer, had been with the CF OA program for over a year; he was posted for duty with the USAFOMC with a reporting date of August. Captain (then Lt) Julia Hoskins, a behavioral scientist working as an Inventory Developer, was selected to represent the USAF in Ottawa, and arrived for duty in National Defence Headquarters in June, after completing Squadron Officers School enroute.

During her two years in Ottawa, Capt Hoskins worked with an operational team, developing job inventories and analyzing results. She visited a variety of CF bases, and gained invaluable experience as an analyst. She briefed several general officers on a number of OA projects and was able to impact on decisions to impliment the recommendations resulting from the OA data analysis. Captain Hoskins was reassigned in August 1983 to the Training Development Service, a

newly created program assigned to the USAFOMC in 1982.

Major Falle initially analyzed an enlisted aircrew specialty and developed a very concise OSR on the results. As a more senior and experienced officer, he was quickly also tasked to help train and guide new civilian and 2Lt analysts in learning CODAP and normal USAF occupational analysis procedures. Major Falle was responsible for an operational trial of the experimental CORSET programs aimed at assisting analysts in verifying significant job group differences (Falle 1983; Phalen 1983). In this project, he developed some alternative indices for comparing job groups; his suggestions were implimented in the final AFHRL CORSET programs as USAFOMC program options

He conducted a number of special projects including the development of a special inventory to evaluate the cross training requirements if three cryptographic and communications equipment repair specialties were merged as proposed. The data collected in this special study were well received and impacted the proposed merger and subsequent training decisions (Knight & Falle 1983). Major Falle was a full participant in all OMC activities. He represented the organization at Utilization & Training workshops held at Lowry and Sheppard Technical Training Centers. He briefed the Hq ATC DCS for Technical Training on OSRs and special projects. He also chaired a paper presentation session at the 24th annual Military Testing Association Conference which was cohosted by AFHRL and USAFOMC in San Antonio in October 1982. Major Falle was reposted to DMOS3 in Ottawa in August 1983.

SECOND WAVE EXCHANGE

In the summer of 1983, the second round of exchange tours was accomplished. Captain (then 1Lt) Randy Agee of the USAF OA program was selected and sent to Ottawa with enroute training at the Squadron Officers School. As an experienced analyst, Capt Agee was experienced in working with training center personnel on the use of OSR data in training decisions. Capt Agee has been assigned to an OA team and recently advanced to being acting team chief for the project, as supervisor of four senior technicians and Warrent Officers. He is currently traveling to operational bases, including some in Europe, managing the development of a comprehensive inventory for a study of civil engineering trades.

Captain Frank Strickland was assigned to the USAFOMC/OMY in August 1983 as the Canadian Forces exchange exchange officer. He is a Personnel Selection Officer (PSO), the CF equivalent of the USAF behavioral scientist. He was stationed at Greenwood CFB in Nova Scotia. He is currently working as an Inventory Developer and will be transitioning to work as an analyst in August of this year.

Under a special arrangement worked out between the CF OA program chief and the CF Applied Research Unit head, a PSO would be sent tdy to Ottawa for six months exposure to the OA program, then for two years to the USAFOMC OA tour. This is programmed to be followed by a two year tour at the National Defence Headquareters (DMOS3) and subsequent assignment to the Applied Research Unit. Thus, a more formal career development program has been planned to move an experienced PSO through the OA area as preparation for a career in Human Resources research.

Lt Col Frank Finch is the present Director Personnel Selection Research & Second Careers and functions as the leader of the PSO branch. He recently expressed interest in expanding the opportunities for

CF-USAF exchange programs. His predecesor was Lt Col Glen Rampton (a PhD psychologist), who retired last year. In a trip report of his visit to Israel to examine how psychologists are used in the Israel Defence Force (IDF), Rampton made a number of significant suggestions on how PSOs could be better utilized in the Canadian military service. His report, submitted in late 1982 just prior to his retirement from active duty, has served a stimulus for a major restudy of the role of the PSO and potential new missions which may considerably expand the work of Canadian military behavioral scientists (Rampton, 1982; see Reference Note 1).

EXCHANGE PROGRAM EVALUATION

To date, the USAF-CF Exchange Program has been an outstanding success. Not only have the primary objectives of the program been met, but there have been a number of unanticipated results. These include the cooperative technical refinement of CORSET in an operational test, the development of a new technology to access cross training requirements of proposed mergers, and the potential development of USAFOMC computer screening programs to quality control input data. The valuable outcomes of these interactions suggest the need for more direct and continuing information exchange capability between the two OA programs.

The establishment of a structured career development program of sequenced assignments for Canadian Forces PSOs is another major unanticipated result. This is one of the few well structured programs for rational, more formalized development of selected individuals by a specified sequence of assignments. This concept has outstanding potential for proper manning of selected key behavioral scientist positions in both the CF and the USAF.

Another implication is that we might find additional exchange programs valuable. Other behavioral scientists and personnel selection officers could benefit from exchange experience. The roles of the PSO in recruiting and selection in the CF are areas which the USAF needs to explore. New exchanges would also provide additional opportunities for an "overseas" (non ZI) tour, which is currently limited for USAF behavioral scientists to either with the Royal Australian Air Force in Melborne (with the RAAF officer being assigned to the Air Force Human Resources Laboratory, Brooks AFB, TX) or the CF exchange in Ottawa. The OMC assignment is the only CF PSO exchange tour available. Presently, both of the existing USAF exchange programs involve occupational analysis, which greatly limits those who are qualified to be selected to participate. The Canadian Forces may want to look at USAF Human Factors, Test Development, Manpower & Personnel Research, and Organizational Development positions as places where CF PSOs could gain valuable experience in an exchange tour.

The existing programs have proved very worthwhile. We all gain from such interactions. They represent a significant investment in the futures of both our people and our programs.

REFERENCE NOTES

1. Rampton, Lt Col Glen. Trip Report of Visit with Israel Defense Force Psychologists. Manuscript report, CF Applied Research Unit, Windsor, ON, 20 Oct 1982.

REFERENCES CITED

Carmark, Bruce. The Canadian Occupational Analysis Program. 1969, in R. E. Christal (Ed), Proceedings of Div 19, Military Psychology symposium: Collecting, analyzing and reporting information describing jobs and occupations. Lackland AFB TX; Occupational Research Division, Air Force Human Resources Laboratory, AFHRL-TR-73-75, 1974.

Christal, R. E. The United States Air Force occupational research project. Lackland AFB, TX; Occupational Research Division, Air Force Human Resources Laboratory, AFHRL-TR-74-19, 1974.

Falle, Ian E. An operational trial of CORSET. in the Proceedings of the 25th Annual Conference of the Military Testing Association, Gulf Shores, AL; Military Testing Association. 27 Oct 1983.

Knight, James R., Jr., and Falle, Ian E. The assessment of common skills among similar career ladders. in the Proceedings of the 25th Annual Conference of the Military Testing Association, Gulf Shores, AL; Military Testing Association. 25 October 1983.

Mitchell, J. L. A short history of job analysis in the military. in the Proceedings of the 25th Annual Conference of the Military Testing Association, Gulf Shores, AL; Military Testing Association. 25 October 1983.

Morsh, J. E. Job analysis in the United States Air Force. Personnel Psychology, 1964, 17, 7-17.

Phalen, William J. CODAP: Recent advances in the identification, evaluation, and description of job types. in Proceedings of the 25th Annual Conference of the Military Testing Association, Gulf Shores, AL; Military Testing Association. 26 October 1983.

Personal and Organizational Factors Affecting Health and Morale of US Army Special Forces Soldiers

Major Frederick J. Manning, USA
Major Terrence D. Fullerton, USA
Walter Reed Army Institute of Research

ABSTRACT

Participant observation, group and individual interviews, and standardized questionnaires were used in assessing the relative levels and major sources of stress for Special Forces "A-team" and support unit soldiers. Team members scored higher than support troops on measures of physical and psychological well-being and satisfaction with their unit, their chain-of-command, and the Army as a whole. Although team members acknowledged conflicts between job and family at double the rate of support unit members, scores of the two groups on the Marital Satisfaction Inventory were not significantly different. Scores for the two groups on measures of sensation seeking, type A behavior and social desirability were also similar, but team members had more "internal" locus-of-control scores and more "protestant ethic" in their values toward work. Reports of social support available from family and friends were similar, but team members reported far more support available from unit and Army. Intense unit cohesion in A-teams is forwarded as the primary source of these findings.

INTRODUCTION

Unit cohesion has become a topic of major interest to the U.S. Army over the past few years. Indeed, a number of major policy and operational changes have already taken place in response to a perceived shortfall in this area, and long-term plans call for the eventual replacement of our individual-based personnel system with a unit-based system similar in many ways to the "regimental" system used by most of our NATO allies (Meyer, 1982).

The importance of strong interpersonal bonds among members of small combat units is unquestionable (see, for example, Marshall, 1978; Mullins & Glass, 1973). Little definitive work backs up claims for its importance in peace-time (however, see Ingraham, 1984; and Manning & Ingraham, in press), despite an extensive but equivocal civilian literature on inter-personal attraction and productivity. The present work adds a small piece of evidence derived from a study of peacetime stress in members of what most observers would judge the Army's most cohesive permanent units - Special Forces "A-teams."

The burgeoning literature on the role of social support (variously defined but generally at least implying strong interpersonal bonds) would suggest that high levels of unit cohesion should provide considerable protection from physical and mental illness (Gottlieb, 1981) as well as high levels of job satisfaction (Cooper & Payne, 1980). Special Forces (SF) soldiers are "triple volunteers" (Army, parachute, SF) however, and as the authors were told numerous times, "a different breed of cat." Fortunately not all such "triple volunteers" end up on "A" detachments, the 12-man operational unit of Special Forces. Much to their dismay, many end up in conventionally organized and operated service, signal, and military intelligence companies providing support to the "A-teams." They also provide important support for this study, by serving to control in some measure for volunteer bias.

METHOD

Unstructured interviews and participant observation both in garrison and on overseas training missions led to the assembly of a large questionnaire composed of the following: Dupuy's (1978) index of General Well Being; Ware's (1979) Health Perceptions Inventory; a 20-item version of the Roach et al., (1981) Marital Satisfaction Scale; Datel's (1978) Army Satisfaction Inventory; Zuckerman's (1979) Sensation Seeking Scale; Bortner's (1969) short measure of type A behavior; Rotter's (1966) locus-of-control scale; an abbreviated version of the Wollack et al. (1971) Survey of Work Values; and original surveys of small unit command climate, attitudes toward and perceived availability of social support, and the frequency of job-family conflicts. Sources of stress were assessed with open-ended questionnaire asking the soldier to list the things he liked most and disliked most about his unit, as well as the things he felt his chain-of-command should start doing and stop doing.

Questionnaires were distributed to two full battalions of "A-team" soldiers and 4 companies of support troops, along with a pre-addressed return envelope. A return rate of approximately 30% yielded samples of 92 and 84 respectively.

RESULTS

Table 1 summarizes the data from the multiple choice portions of the questionnaire. The maximum range figures are set up so that the rightmost figure represents the "best" possible score in the case of the state measures, and a high level of the measured trait in the case of the trait measures. It is thus apparent that A-team soldiers were more satisfied with their General Well-Being (primarily a mental health index), their physical health, the Army as a whole, and the command climate in their own unit. Both groups reported high and statistically similar levels of satisfaction with their marriages, but team members reported job-family conflicts twice as often as support unit soldiers. Relevant here are the results of an instrument inadvertently included only in the A-team questionnaires. This instrument asked respondents to indicate on a 5-point scale the stress caused them in the prior month by each of 26 potential stressors. "Competition between SF and family" ranked 23rd, with only career security, financial difficulties, and long hours rated as less stressful.

Personality differences were seen only on the Locus-of-Control measure, where team members appear more "internal" (i.e. tend to see a person's circumstances controlled more by his or her own actions than by luck, fate, powerful others, or other causes external to the person), and in the two Survey of Work Values factors reflecting job involvement & pride in work.

Analysis of answers to the open-ended questions on what the soldier liked and disliked about his unit and chain-of-command showed primarily differences in degree rather than in areas of concern. The most frequent "like" of each group was some aspect of unit cohesion (though #2 for team members was the intelligence, competence or motivation of their teammates, while the support units' second choice was travel and time off). Both groups listed lack of realistic training as their primary "dislike," and both faulted their chain of command for making decisions affecting them without input or feedback from below. In addition, A-team members found the number and type of support details required of them particularly galling, while support unit soldiers more frequently cited both undesirable personal traits in peers and unequal treatment by the chain-of-command based on rank, race, age, job specialty, or some other personal characteristic.

TABLE 1: Mean Values of SF Soldiers on Questionnaire Variables

Variable (maximum range)		A-Teams	SPT Units	p[a]
States				
General Well-Being Scale (0-120)		84.7	73.3	.0001
Health Perceptions Inventory (160-32)		69.2	74.7	.004
Marital Satisfaction Scale (0-100)		83.0	86.8	NS
Army Satisfaction Inventory (60-300)		207.1	194.4	.006
Command Climate Survey (0-13)		8.8	6.8	.0001
Job-Family Conflict (2-0)		1.8	0.9	.0001
Soldier Social Spt Scale:	Family (0-10)	5.1	5.4	NS
	Friends (0-10)	6.2	5.9	NS
	Unit (0-10)	7.8	5.6	.0001
	Army (0-10)	8.1	6.4	.0001
Traits				
Sensation-Seeking Scale (0-40)		20.0	20.9	NS
	Thrill Seeking (0-10)	8.7	8.4	NS
	Experience Seeking (0-10)	4.3	4.5	NS
	Disinhibition (0-10)	4.5	4.8	NS
	Boredom Suscept (0-10)	2.6	3.0	NS
Bortner Type A Scale (14-70)		43.0	44.2	NS
Rotter Locus-of-Control (0-23)		8.7	10.1	.008
Crowne-Marlowe Social Desirability (0-38)		19.5	19.5	NS
Work Values:	Intrinsic worth (5-15)	13.5	11.9	.0001
	Organization ethic (5-15)	13.6	12.5	.002
	Upward striving (5-15)	11.1	10.9	NS
	Social status (5-15)	9.6	9.7	NS
	Conventional ethic (5-15)	13.2	12.5	NS
	Attitude to earnings (5-15)	7.7	8.1	NS
Soldier Social Spt Scale:				
	Pos Orient (0-10)	6.6	6.8	NS
	Provides Spt (0-10)	8.4	7.8	NS

[a] Group differences tested in all cases with the General Linear Models Procedure of SAS Institutes' Statistical Analysis System.

Finally, as Table 2 shows, the two groups differed substantially in demographics: team members were older, more educated, higher ranking, had been in both the Army and SF longer, were more often married and less often from a minority ethnic group.

TABLE 2: Demographic Description of "A-Team" and Support Unit Samples

Variable	A-Team	Support Unit
Median Age	29	23.5
% Caucasian	90	62
% Educated beyond HS	68	43
% E1 to E4	3	46
Leading Job Fields	Infantryman Combat Engineer Combat Medic	Intelligence Supply Administration
% Married	92	29
% Divorced or Separated	1	9
Median Months in Army	108	41
Median Months in SF	52	25

DISCUSSION

As predicted by social support theory, the members of the highly cohesive 12-man A-team reported higher levels of physical and psychological well-being and greater satisfaction with their unit, their chain-of-command, and the Army as a whole. Not surprisingly, they also perceived their unit and the Army as sources of help in times of personal difficulty far more often than SF soldiers assigned to support units.

As expected, the support unit soldiers served as excellent controls for the "triple volunteer" personality, since they differed from team members on only 3 of 16 measured traits. It might even be argued that these 3 "traits" are in fact not immutable, and that the group differences may in fact be a result of group membership (Nowicki & Barnes, 1973). Team members for example, clearly have more control over their daily activities than support troops, so it would be remarkable if they did not show up as more "internal" on the locus-of-control scale.

The support unit soldiers were however decidedly different from the team members in background demographic characteristics. We are in the process of addressing this difficulty by extending the study to include samples from non-SF infantry and airborne units. Until such time as that analysis is possible we would offer the present findings as strong support for the desirability of high levels of unit cohesion in peacetime as well as in combat.

REFERENCES

Bortner, R. W. A short rating scale as a potential measure of pattern A behavior. J. Chron. Dis., 1969, 22, 87-91.

Cooper, C. L. and Payne, R. (Eds). Current Concerns in Occupational Stress. London: Wiley & Sons, 1980.

Crowne, D. P. and Marlow, D. A. A new scale of social desirability independent of psychopathology. J. Consult. Psychol., 1960, 24, 349-354.

Datel, W. E. Final Evaluation Report on Fort Ord Project Volar. Alexandria, Va.: Defense Technical Information Center, 1978.

Dupuy, H. J. Self-representations of general well-being of American adults. Paper presented at meeting of American Public Health Assoc., Los Angeles, 1978.

Gottlieb, B. H. Social Networks and Social Support. Beverly Hills: Sage, 1981.

Ingraham, L. H. The Boys in the Barracks. Phila: Institute for the Study of Human Issues, 1984.

Manning, F. J. and Ingraham, L. H. "An Investigation into the Value of Unit Cohesion in Peacetime." In Belenky, G. L. (Ed) Contemporary Studies in Combat Psychiatry. Boston: Kluwer, in press.

Marshall, S. L. A. Men Against Fire. Gloucester, Mass.: Peter Smith, 1978.

Meyer, E. C. The Unit. Defense 82, 1982 (Feb), 2-9.

Mullins, W. W. and Glass, A. J. Neuropsychiatry in World War II, vol. 2, Overseas Theaters. Washington: Office of the Surgeon General, Department of the Army, 1973.

Nowicki, S. and Barnes, J. Effect of a structured camp experience on locus of control orientation. J. Genetic Psychol., 1973, 122, 247-252.

Roach, A. J., Frazier, L. P. and Bowden, S. R. The Marital satisfaction scale: development of a measure for intervention research. J. Marr. & Family, 1981, 43, 537-546.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psych. Monog.,1966, 80, Whole No. 609.

Ware, J. E., Johnson, S., Davies-Avery, A. and Brook, R. Conceptualization and Measurement of Health in Adults in the Health Insurance Study: vol. 3, Mental Health. Santa Monica: Rand Corp. 1979.

Wollack, S., Goodale, J. G., Wijting, J. P. and Smith, P. C. Development of the survey of work values. J. Applied Psychol., 1971, 55, 331-338.

Zuckerman, M. Sensation Seeking: Beyond the Optimal Level of Arousal. New York: Plenum, 1979.

DISCLAIMER

WORKSHOPS FOR MANAGEMENT OF COMBAT STRESS

A. David Mangelsdorff, Ph.D.,M.P.H.
James M. King, Ph.D.
Donald E. O'Brien, Ph.D.

Health Care Studies and Clinical Investigation Activity
United States Army, Health Services Command
Fort Sam Houston, Texas 78234

Abstract

Combat produces stresses which may result in psychiatric casualties. Commanders need to be aware of how to assess the psychological readiness of their units and of the means to develop cohesion. The results of three workshops on combat stress and related issues are reviewed. In addition, future directions are discussed.

Much has been written about how combat produces stress and the possible levels of psychological casualties in combat. Commanders need to be reminded of the threat of psychological casualties and how to prepare for dealing with these casualties. In addition, commanders need to be able to assess the readiness of their units, particularly in terms of morale and cohesion.

There have been a variety of programs developed to help manage stress. At the Eighth Symposium of the Psychology in the Department of Defense (1982), a paper session on stress management documented the variety of civilian and military efforts being undertaken to deal with stress (Swiney, 1982). Stress programs may focus on individual responses, group responses, organizational responses, situational factors, or some combination of the factors. Our Army program was developed in response to repeated requests for information on how to develop training programs for the management of combat stress reactions; there have been three workshops conducted thus far. The purpose of this presentation is to review what has been done and to suggest some future directions for our Army program.

First Users' Workshop on Combat Stress.

The first Users' Workshop on Combat Stress attempted to address the needs of the mental health care providers in several key Army combat units (the 82nd Airborne Division, the 101st Airborne Division, the 2nd Armor Division, and the 1st Cavalry Division). These needs included: (1) presenting authoritative information on current threat estimates, concepts on countering and [illegible] the threat, and casualty estimates; (2) discussing the [illegible] tasks and functions of line and health care personnel

who will identify, refer, treat, and/or prevent combat stress casualties; and (3) setting goals, establishing methodologies to achieve these goals, and deciding upon means for evaluating goal attainment.

Participants from the combat units were asked to bring and to describe whatever training programs, handouts, packets, or written ideas they had for training soldiers, leaders, medical, and mental health personnel. The mental health staffs were to be prepared to: identify their unique training needs, commit themselves to developing and conducting their own training programs, evaluate their own programs, and share the results of their programs and evaluations with the other workshop participants.

Three tasks groups were formed; groups were instructed to define their goals, decide how to reach the goals, and determine how to evaluate the progress toward achieving those goals. Most goals focused on establishing training programs and reorganizing resources to achieve maximum effects.

A training program developed for the Community Mental Health Activity at Fort Knox, Kentucky, entitled Project COPE (Combat Operations and Psychiatric Effectiveness) was run for the workshop participants. The three task groups were dissolved into three mixed groups which were required to role play medical personnel at Battalion Aid Stations. Participants were required to triage psychiatric and medical casualties (represented by analog field medical cards), provide effective interventions with soldiers presenting stress reactions, maintain effective radio communications, and cope with increasingly stressful demands. Evaluators provided feedback to participants on the effectiveness of their treatment and dispositions. After the exercise, the experiences of the participants were processed for feelings and insights into how a training program might be developed.

Participants were asked to work toward achieving the goals defined in their task groups. A network of resources had been established. Contributions from the participants were collected, edited, and assembled into a proceedings which was sent to each participant (Mangelsdorff and Furukawa, 1981). The collected proceedings was intended to serve several functions, as: a reference resource, a commitment toward disseminating information, and a reminder of intended goals and proposed methodologies. It became clear that additional workshops were needed to reach other Army units; this lead to the second Users Workshop on Combat Stress.

Second Users' Workshop on Combat Stress.

The Second Users' Workshop brought together both line officers with command or training responsibilities and mental health officers. Participants were asked to be willing to exchange their own training materials. This diverse group

attended the Second Users' Workshop, which allowed for modifications in the manner and type of presentations.

Three task groups were formed. Groups were asked to assess the needs of the members in terms of concerns, problems, or issues related to combat stress. Solutions for the needs were to be developed. Organizational Effectiveness consultants were used to facilitate the task group process. The most common themes were: development and presentation of an effective combat stress program, determination of who needed the program, and determination of where the program was needed.

Participants were tasked with returning to their respective posts to work on developing effective training programs. As in the First Users' Workshop, contributions from the participants were assembled into a proceedings (Mangelsdorff and Furukawa, 1982). The proceedings from the Second Users' Workshop were sent to all participants in both the First and Second Users' Workshops. The intent was to remind participants of their commitments and to enlarge the network of individuals working on the problems of combat stress.

Third Users' Workshop on Combat Stress.

The Third Users' Workshop was devoted to unit cohesion, a crucial determinant of individual and unit psychological readiness. Representatives from the 4th Infantry Division, Walter Reed Army Institute of Research, the Soldier Support Center, the Army Research Institute for the Behavioral and Social Sciences, the 9th Infantry Division, the 82nd Airborne Division, the Academy of Health Sciences, and the Israeli Defense Force met to discuss what they were doing with respect to the assessment and development of cohesion. The participants were tasked to: define the elements of cohesion, determine what commanders need to know about the cohesiveness of their units, identify indicators and/or crucial aspects of unit cohesion, determine how best to provide feedback to commanders about the cohesiveness of their units, and develop suggestions to assist in the development of unit cohesion.

Groups were formed to meet the tasks. Facilitators from the Health Care Studies and Clinical Investigation Activity assisted the process. The results from the separate groups were presented to all participants for further discussion and reflection. The definition and assessment of the elements of cohesion revealed a multi-faceted process.

As in the previous workshops, contributions from the Third Users' Workshop participants were collected and assembled into a proceedings (Mangelsdorff, King, and O'Brien, 1983). The proceedings were distributed to all participants for further consideration.

It was apparent that another workshop would be required to focus on the lessons learned from the recent Israeli experiences. This involved collecting the papers and presentations made at the various conferences on stress and combat mental health. The Fourth Workshop will bring together participants from the Academy of Health Sciences, Walter Reed Army Institute of Research, Health Services Command, and the Israeli Defense Force to examine what lessons can be derived from the Israeli involvement in Lebanon and in the 1973 wars. Differences in organizational and cultural factors will be considered. This fourth workshop is tentatively scheduled for late summer in 1984.

The purpose of the workshops has been to bring together as many individuals as possible who are engaged in research, teaching, or consulting in the area of combat stress. The workshop model allows participants to learn the current trends and to interact with other professionals. As the network of participants expands, it is hoped that more combat and combat support commanders will appreciate the need to prepare and train for dealing with combat stress reactions.

References

Swiney, J.F. (Editor) Proceedings, Psychology in the Department of Defense, Eighth Annual Symposium. USAFA TR 82-10. 1982.

Mangelsdorff, A.D. and Furukawa, T.P. (Editors) Proceedings Users' Workshop on Combat Stress. Academy of Health Sciences, 1981.

Mangelsdorff, A.D. and Furukawa, T.P. (Editors) Proceedings Second Users'Workshop on Combat Stress. Academy of Health Sciences, 1982.

Mangelsdorff, A.D., King, J.M., and O'Brien, D.E. (Editors) Proceedings Third Users'Workshop on Combat Stress. Cohesion. Health Care Studies and Clinical Investigation Activity, 1983.

Combat Psychiatry
A Primary Prevention Program in the 4th Infantry Division (Mechanized)

LTC Linton S. Holsenbeck, M.D.

Abstract

The 4th Infantry Division (Mechanized) Mental Health Section has developed a combat psychiatry program consisting of four elements: 1) Battle Stess Training 2) the Combat Stress Survey 3) Battlefield Interviewing 4) Brigade Consultation Teams. This program is presented as a model for combat-oriented preventive psychiatry. Its evaluation and initial impact are outlined.

Background

The first intentional assignment of a psychiatrist to a combat division in 1943 marked the beginning of the modern era of combat psychiatry. This event was no innovation but rather a re-embracing of principles established by Thomas Salmon in WWI. The immediate stimulus for the measure was the unexpected incidence of combat stress casualties in North Africa. These casualties were a surprise in part because of a mistaken belief that men predisposed to stress casualty status could be recognized at induction and kept out of the combat environment. This overly optimistic screening effort was the first major attempt to prevent combat stress casualties on an organizational basis.

While screening largely failed, alert mental health professionals assigned to combat divisions quickly began to work backward from the casualty to his military unit. There they began to recognize organizational precursors to casualty status, to develop behavioral concepts which had parallels in leadership, and to provide useful advice to commanders. The crux of this development was the marrying of behavioral to leadership principles. Through the intercourse of this marriage the conclusions of observors with an entirely different perspective on the fighting man, observors like Clausewitz and S.L.A. Marshall, were comingled with the observations of behavioral scientists. The offspring was a body of knowledge with great potential usefulness for both leaders and behavioral scientists. This benefit was fully realized in the Korean conflict when combat psychiatric principles were competently applied from the beginning.

So what happened? Within a decade military psychiatry seems to have demilitarized. As much as anything else, this probably resulted from the turnover of personnel, younger mental health professionals having been schooled in civilian community psychiatry. Social service programs of all sorts burgeoned in the Army. Division Mental Hygiene Sections became "Human Resource Centers" and vacated their combat role. The marriage between leader and behavioral scientist that had appeared so full of promise in the early Fifties had deteriorated to reluctant cohabitation by the late Sixties and continued, with the exception of occasional threats of divorce, through the Vietnam era.

In Vietnam air supremacy allowed centralized treatment, "proximity" came to mean anywhere inside the country, and substance abusers and bad actors went

unrecognized as combat stress casualties until a decade later when they were reenfranchized as "Post-traumatic Stress Disorder".

Only in the last five years has the word cohesion, spoken with a slight Hebrew accent, returned to our vocabulary. Thirty years late, S.L.A. Marshall is the darling of the combat arms. One finds the phrase "Personnel Management" lined through in old lesson plans and the term "Leadership" discretely pencilled in its place. "Combat Stress" and "Cohesion Technology" symposia attract the same participants. It seems time for the rebirth of Combat Psychiatry.

The Program

In the fall of 1982 the leadership of the Mental Health Section did a self-assessment. Serious flaws were noted in our method of operating and in the product of our work. It wasn't that things were being done badly, actually everything was being done quite well. With the hospital Community Mental Health Activity we ran an efficient hospital-based outpatient clinic, seeing lots of patients.

So what was wrong? As a Division Mental Health Section our mission, since 1943, was to serve the psychological readiness of the Ironhorsemen, to train and maintain in the prevention, recognition, and management of combat stress casualties, and to be combat ready and prepared to mobilize. If that was us, then we no longer existed. Our organization bore no resemblance to that definition. But since we knew of no other MHS which resembled that definition either, why change?

That question was never answered in more than philosophical terms but change we did and the results have been extremely positive. The MHS is now firmly identified with the Ironhorse Team and is no longer an orphaned cousin of the hospital. Our own morale, cohesion, and sense of purpose have improved significantly.

We began by donning the BDU's, establishing a chain of command internal to the MHS, sending everyone downrange at the first opportunity, and making the Battalion Run mandatory for MHS personnel. A Battle Stress Training module was developed, MHS personnel were trained, and in turn trained their medical companies. The training was then offered to the division. During 1983 over 500 officers and 350 NCO's were trained, usually as a topic in ongoing Officer and NCO Professional Development seminars. Chaplains and medical personnel were given special attention and a block was included in the precommand LEAD course for company grade officers. Battle Stress Training achieved the initial exposure and credibility the MHS needed to be recognized as Ironhorsemen and laid the foundation for the Combat Stress Survey.

We had been intrigued by the mysterious "Israeli Morale Survey" and even obtained a copy, which later turned out to be bogus. Using our own knowledge of the combat stress literature we designed a survey instrument for the purpose of measuring psychological readiness. We specifically intended the survey to provide only a snapshot, to be brief and combat relevant, and to not exceed the proverbial twenty questions. By mid-year the instrument itself, our method of conducting it, and computer assisted analysis had jelled. We

were ready for a major study. The 1st Brigade Task Force Study was accomplished around the brigade's October deployment to the National Train Center at Fort Irwin, California. It will be presented in detail later in this paper.

The third component of our program was also developed during the October NTC deployment. The reunion between behavioral scientist and military leader had to be accomplished. We selected an armored battalion within which to attempt battlefield consultation. Our task included not only the challenge of providing accurate, useful consultation to commanders in mock combat but also survival and safe movement in a fluid, rigorous, and mildly hostile environment. A brief, structured interview, the Battlefield Interview was drafted and enlisted behavioral science specialists were instructed in rough sampling of deployed soldiers. By using intrinsic transportation, we were able to move back and forth from field trains to combat trains to company trains to fighting positions, from brigade headquarters to maneuver battalions to support battalion. We interviewed during lulls in fighting, withdrawing to the relative safety either of the battalion rear or of "thick-skinned" vehicles during periods of fighting.

We could readily detect level of morale and could often localize sources of poor morale. Examples included a delay in mail delivery, a badly managed battle, and a company commander who had somehow incurred a passive rebellion. Our observations on morale, cohesion, and fighting spirit correlated very closely with performance in battle and with the observations and intuition of battalion and brigade commanders. We seldom suprised them but rather confirmed their own assessments with semi-objective data unavailable to them by other means. Our consultation was deemed extremely useful by the commanders to whom we consulted and we observed leadership actions taken, in part, based on our recommendations.

The most recent addition is a consultation program with MHS staff assigned, both in the clinic and in the field, to a designated brigade. In addition to seeing referrals from their brigades, Brigade Consultation Team counselors will develop statistical data traditionally indicative of organizational climate. through an orientation program they will spend time with the various units in their brigades in the garrison workplace, in the barracks, and in the field. They will serve as observors, sensors, and conduits of information for the Brigade Consultant who will be a designated behavioral science officer. This will round out our program and provide us the basic armamentarium for our combat mission. From this base we can develop more specialized and refined means for managing specific problems and providing specific services.

1st Brigade Task Force Study

This study is presented because of its key importance in the evolution of our relationship with the division. As stated earlier, by mid-1983 we had completed enough preliminaries to believe that the Combat Stress Survey was capable of meeting the following objectives:

1. to be quick and easy to administer and take.
2. to yield an accurate though general assessment of the soldiers' sense of cohesion in teams, confidence in leaders, confidence in individual

fighting skills, and perception of unit effectiveness.

3. through successive administrations, to define the impact of intercurrent experience on psychological readiness.

4. to yield internal standards against which to compare units.

The units studied included an infantry company, two armor companies, the brigade headquarters company, a medical company, a signal platoon, a composite military intelligence company, and a chemical company composed of a platoon each from the two division chemical companies. This diversity had advantages in that we learned early that many definitions of "team" exist in different types of units, that a "unit" may not be a unit in any sense except the sharing of a guidon and crest, that many "commanders" command little outside their orderly room, that six soldiers from the same platoon may have six different faces in mind when they respond to an item about "my leader", and that outside the combat arms there exist many units in which the designated chain of command can have little impact on morale and cohesion.

The disadvantages were that the diversity diluted the impact of our findings. Unless "units" are defined in highly individualized ways, we have difficulty establishing correlations between the actions of leadership and changes in morale. Without these correlations our data is interesting but useless.

Our study confirmed that the CSS is quick and easy to give and to take. This was proven upon our discovery that the entire stock of questionaires, answer sheets, and pencils brought to Ft Irwin for the redeployment survey had been lost. The brigade adjutant was able to make 500 copies from a single dusty extra recovered from the brigade commander's briefing folder. We scrounged pencils from Weed Army Hospital and forewent the answer sheets. Units were surveyed in a large, utterly barren, outdoor staging area while awaiting transportation to the airport. Administration, even under these conditions, took only about twenty minutes.

Our second hypothesis was also confirmed. The CSS yielded a profile of each company which correlated well with the impressions of battalion and brigade commanders and with battlefield interviews and informal observations by MHS personnel. The survey yielded measurements of the four general parameters which interested us and a diversity of company profiles which seemed to "fit" when briefed to the commanders of those companies.

The NTC experience was clearly demonstrated to have to have measurable impact on companies that were studied. That impact is a subject of great interest to commanders and has developed into a study in and of itself which will be conducted through successive brigade deployments over the next year.

Finally, the CSS yielded comparisons which were of intense interest to commanders. We were able to show commanders how they compared to themselves over time, to other companies, and to the task force as a whole.

We believed CSS results would be most useful to more junior commanders. Officers successful enough to land battalion and brigade commands generally hav a good "feel" for psychological readiness of their soldiers but at the company grade level this "feel" is just being developed. An external measure against which to compare one's own estimation and from which to make adjustments would seem extremely useful. To enhance this effect we asked company

commanders to estimate their unit's responses and to select items relevant to them and establish goals prior to the feedback session. For example, on the item "I am proud of my unit" a commander might estimate that 60% of his unit would answer positively. His goal might be 90%. If his actual result was 40% he would be forced to examine his means of estimating pride among his soldiers and would know he had some way to go to achieve his goal. We also asked commanders to estimate their strongest and weakest items and to pick out the items they felt were most affected by the NTC experience.

The actual feedback session covered a rank ordering of items, average scores for each item, significant changes between surveys, individual company performance plotted against various standards internal to the brigade, an item by item analysis by the percentage of respondents answering each item positively and negatively, and an analysis by demographic groups. Feedback sessions were held with the company commander and those of his choice. Feedback sessions generated a great deal of discussion and thinking about the problems of cohesion, morale and confidence. They also generated requests for resurvey at times that commanders felt would be useful, like before a change of command and during yearly training evaluation.

The only outcome about which the MHS staff is ambivalent is the tremendous demand for further studies generated by this initial success. At the time of this writing a much larger study is underway which will assess the entire Division Support Command. Another NTC-centered study has been contracted with the 3rd Brigade Task Force which deploys in January 1984. This study will focus on two entire maneuver battalions, one infantry and one armored.

Summary

In eighteen months the 4th Infantry Division Mental Health Section has transitioned from a traditional hospital-bound outpatient clinic to an effective combat psychiatric team engaged in a broad range of activities to immunize division fighting units against inordinate combat stress casualty rates. While we hope a true test of our effectiveness is never made, we are confident that our approach is well chosen. Initial reaction of leaders with whom we work is supportive. Their reception has been positive and resistance has been remarkably absent.

Divisions never know what to expect of mental health people anyway. That's why it is so easy to withdraw into the safety of our clinics and ignore our combat mission. The other thing about divisions is that once we make ourselves visible, we're expected to deliver tangible results. This reaction more than any other seems to frighten mental health professionals. It's part of the reason we're not surgeons or engineers. The hard fact is that modern warfare being what it is, nearly every behavioral scientist now in uniform will have a combat role in any major conflict. Best we prepare now.

Acknowledgements

The author acknowledges the work of CPT Brad Powell, Ph,D and CPT Liz Donald in development of the Combat Stress Survey and of SGT Chris Wagoner for his unflagging support in overcoming the inertia in our own organization. Outline of Battle Stress Training, copies of Combat Stress Survey and Battlefield Interviews are available on request from the author (Autovon 691-5522).

AD-P003 285

IDENTIFICATION, MEASUREMENT AND PREDICTION OF A PSYCHOTIC VARIABLE WITH NEURO-PSYCHOLOGICAL MEASURES.

Vincent A. Escandell, Ph.D.
USAF Rgn Hospital Sheppard, Sheppard AFB, TX

The cognitive aspects of functional disorders were identified and measured with an ex-post facto design wherein twenty psychotic subjects, twenty non-psychotic psychiatric subjects and twenty normal subjects, matched for age, sex, education, and race, were tested with fifteen organically validated cognitive test variables. The cognitive tasks were grouped and found significant by MANOVA designs in the categories of attention concentration tasks, right hemisphere or spatial tasks, and anterior oriented tasks. The task categorizations not found significant were the left hemisphere or language oriented tasks and posterior oriented tasks. The right anterior oriented tasks were specific to separating the psychotic group from other groups. A regression prediction for severity of psychopathology was even more unique to the right anterior tasks.

This prediction equation was cross validated with twenty psychotics. The predictor validity coefficient was .52. Then for 82 patients the predictive equation using a base rate of .24% had a selection ratio of 58%, a selection rate of 90% and a utility rate of 33%. A comparison of correct selection vs the MMPI 8 scale is presented.

The literature is replete with studies wherein psychotic groups display a decrement of a general nature on cognitive tasks. Psychiatric populations in relation to level of performance on cognitive tasks have followed this sequence: organic < schizophrenic < affective psychotic < neurotic < personality disorder (Heaton, Baade & Johnson, 1978); Brilliant & Gynther, 1963; Overall, Hoffman & Levin, 1978; Miller, 1975). Lilliston (1973) concluded the number of tasks failed rather than specific task performance is positively related to the degree of functional impairment. Prior studies used cognitive tasks that were not organically validated or tasks that required predominantly motor involvement rather than information processing or tasks that were focal or singular in function rather than global. Therefore the purpose of this study is to identify, measure and predict the performance of psychotic groups on brain behavior validated cognitive instruments that review focal and global functions.

The primary investigation of the cognitive aspects of functional disorders provided an ex post facto design where twenty psychotic psychiatric subjects, twenty non-psychotic subjects and twenty normal subjects were tested with eleven brain sensitive cognitive based tests. The eleven tests were: the 1) arithmetic and 2) digit span tests from the WAIS-R; the 3) trails A & 4) B from the Reitan Halstead adaption of the Army Alpha subtest; the 5) Temporal Orientation test, 6) Three Dimensional Constructional Praxis test, 7) Judgment of Line Orientation test, 8) Controlled Word Association test, 9) Token Test, 10) Facial Recognition Test and 11) Digit Sequence Learning Test from Benton's Contributions to Neuropsychological Assessment. The predictive validity equation study used twenty psychotics who consecutively appeared at an AF regional hospital from the Air Evac system. The comparison study of the MMPI and cognitive code was based on a selection of all patients (N=82) who took these tests at

Sheppard AFB from 1982-1983. These tests were all individually validated with organic groups. Psychotic subjects were defined as subjects who were actively hallucinating. The subjects were matched by sex, race (white), mean age (34.0), and mean education (12.9 years). Two major questions were investigated by the primary study: 1) This study examined the general notion that there were no significant differences between psychotic psychiatric, non-psychotic psychiatric, and normal groups on cognitive neuropsychological tests. 2) This study investigated questions specific to cortical geographic entities or types of cognitive functioning that differentiated the psychotic group from other groups. The predictive validity study determined the capacity of this regression equation to improve diagnosis. The purpose of the comparison study was to determine this equation's usefulness when the equation was compared to a commonly used classification system designed to discriminate a psychotic group.

The cognitive test scores were analyzed by ANOVA and MANOVA designs with the independent variable being the severity of the functional disorder. Eleven of the fifteen task performances were significant ($p < .05$) in differentiating among the three groups in ANOVA designs. The mean number of tests failed was psychotic > non-psychotic > normal. Three t tests using mean number of test failures showed significant differences between each possible pair of groups. The cognitive tasks were grouped and found significant by MANOVA designs in the categories of attention concentration tasks, cognitive speed tasks, memory and learning tasks, right hemisphere or spatial tasks, and frontal lobe tasks. The task categorizations not found significant were the left hemisphere or language oriented tasks and the posterior oriented tasks.

The results of this study were significant when reviewed for specific geographic and cognitive functions which separate psychotic groups from non-psychotic psychiatric or normal groups. In support of Valenstein & Heilman's (1979) and Damasio's (1979) hypotheses of cerebral mechanisms of emotion, the right frontal lobe tasks were specific to separating the psychotic group from other groups. These tasks included the Temporal Orientation Test (TOT), the Benton Visual Retention Test (BVRT) (scored for correct performances and errors), and the Trail Making Test - form A and form B (TMTA), the regression equation for the severity of psychopathology was:

severity of psychopathology = 1.49+.13307 BVRT - .09165 TOT - .00849 TMTA. Wherein < 2 = psychosis.

The tests selected as best predictors are even more unique to the right frontal lobe. Selected tasks as predictors were: the Benton Visual Retention Test, scored for the number of correct performances; the Trail Making Test, form A; and the Temporal Orientation Test. Therefore, in order to integrate this study's findings with the similarities and diversities of Luria (1966), Reitan (1959), Valenstein & Heilman (1979), and Damasio (1979), this investigation concludes that:

1) right frontal lobe tasks are the most complex tasks for the human repetoire, or 2) right frontal lobe tasks are the highest level behavior of any functional system's hierarchy, or 3) the contribution of the right frontal lobe, being at the highest level of complexity and integration of various cognitive functions makes the performance of tasks which are sensitive to lesions of the right frontal also vulnerable to lesions elsewhere in the brain that may compromise the performance of a component of a task at a lower level of complexity.

In the functional disorders one does not see the type of changes in cognitive abilities that one sees in focal or diffuse brain damaged patients such as the syndromes of amnesia, dementia, aphasia, apraxia and agnosia. However,

changes in cognition reflecting reduced efficiency were observed in the functional disorders and on this basis it is reasonable to speculate that the effect of functional disorders is to disrupt the organization of cognitive abilities at their highest level of integration rather than affecting the primary abilities.

As a measure of its applicability to Air Force medical treatment, the prediction equation was cross-validated with twenty psychotics of varied sex, age, race, education and diagnosis. The predictor validity coefficient was a significant .52.

Reviewing the psychometric sheets of all 82 patients given this battery since 1 February 1982, the predictive equation using a base rate of .24 had a selection ratio of 56%, a selection rate of 90% and a utility for increase in prediction of 33%.

This can be compared to the Schizophrenia Scale (8) of the Minnesota Multiphasic Personality Inventory. This scale was empirically derived to reflect the psychotic pattern of Schizophrenia (Rodgers, 1972, Dahlstrom and Walsh, 1965). Benton (1945), Rodgers (1972) and others vary the rates of selecting schizophrenics from 20-50%. Of importance in this matchup is Gilberstadt and Duker's (1965) finding that of the nineteen known actuarial profile types the 8 scale was always in the profile type except for the infrequent addition of the spike 9 scale. The 8 scale showed greatest relationship to the frequency of checked items for auditory hallucinations, schizoid behavior, depersonalization, ideas of reference and persecution. This scale on the same 82 subjects given the neuropsychological tests could only produce a selection ratio of 16%, a selection rate of 30% and utility rate of -.08.

The equation rather than the 8 scale appears to best reflect the level of cognitive disruption elicited by the presence of psychosis. This equation also suggests a geographic localization which is associated with psychotic disruption. This equation reflects the degree of severity of functional impairment. The significance of the study is best revealed by the superior ability of this equation alone to predict psychosis over self report inventories or other non-neuropsychological measures to identify the presence of psychosis.

TABLE I

MEMBERSHIP OF 82 SUBJECTS IN MMPI VS COGNITIVE EQUATION STUDY

Groups	n	Means
Psychotic	20	1.82
Organic	25	2.03
Nonpsychotic Psychiatric	34	2.26
School failure	2	2.20
Legal	1	2.10

TABLE 2

CLASSIFICATION OF SELECTIONS AS PSYCHOTIC BY CRITERIA VALUES

	Psychotic		Non-Psychotic	
	True Positive	False Negative	False Positive	True Negative
MMPI (Scale 8 > 70)	6	14	32	30
Cognitive Equation (< 2)	18	2	14	48

In summary this study based on neuropsychological tests suggests a geographic location and a specific equation for the presence of psychosis. This equation was shown to be a better predictor than existing objective measures for psychosis.

Selected References

Benton, A.L., Hamsher K. deS., Varney N.R., & Spreen O. (1983), Contributions to Neuropsychological Assessment. New York: Oxford.

Benton, A.L. (1968), Differential behavioral effects in frontal lobedisease. Neuropsychologia, 6, 53-60.

Benton, A.L. (1945), Minnesota Multiphasic Personality Inventory, Revised Edition. In O. Buros (Ed.) The Third Mental Measurements Yearbook (pp. 104-106).

Brilliant, P.J. & Gynther, M.D. (1963), Relationship between performance on three tests for organicity and selected patient variables. Journal of Consulting Psychology, 27, 474-479.

Dahlstrom, W.G. & Welsh, G.S. (1965), An MMPI Handbook. Minneapolis: The University of Minnesota Press.

Damasio, A.R. (1979), The Frontal Lobes. In K.M. Heilman and E. Valenstein (Ed.), Clinical Neuropsychology. (pp 360-412). New York: Oxford University Press.

Gilberstadt, H. & Duker, J. (1965), A Handbook for Clinical and Actuarial MMPI Interpretation. Philadelphia: Saunders.

Heaton, P.K. Baade, L.E. & Johnson, K.L. (1978) Neuropsychological test results associated with psychiatric disorders in adults. Psychological Bulletin, 85, 141-162.

Heilman, K. & Valenstein E. (1979), Clinical Neuropsychology. New York: Oxford University Press.

Lilliston, L. (1973), Schizophrenic symptomatology as a function of probability of cerebral damage. Psychological Bulletin, 82, 377-381.

Miller, W.R. (1975), Psychological defeci' in depression Psychological Bulletin, 46, 1315-1322.

Overall, J.E., Hoffman, M.G. & Levin H. (1978), Effects of aging, organicity, alcoholism and functional psychopathology on WAIS subtest profiles. Journal of Consulting and Clinical Psychology, 46, 1315-1322.

Reitan, R.M. (1959), The comparative effects of brain damage on the Halstead Impairment Index and the Wechsler-Bellvue Scale. Journal of Clinical Pathology, 15, 281-285.

Rodgers, D.A. (1972), Minnesota Multiphasic Personality Inventory. In O.K. Buros (Ed), The Seventh Mental Measurements Yearbook (pp 243-250). New Jersey: Gryphon Press.

PSYCHOLOGICAL EVALUATION OF AIRCREW MEMBERS

John C. Patterson, Ph.D.
USAF School of Aerospace Medicine
Brooks AFB, Texas

Rated USAF Officers requiring psychological evaluation for fitness to fly are referred to the USAF School of Aerospace Medicine (SAM), Clinical Neuropsychology Function (NGN). Other referrals to NGN result from requests for consultation from medical branches within the consultation service. This paper will discuss the approach and psychological methods used in making clinical and aeromedical decisions in the USAF. Further, criticisms of the current procedures will be discussed as well as suggestions for clinical research neede to improve psychology's participation in the aeromedical evaluation of aircrew members.

An aviator may be referred to SAM for a variety of reasons but two features are always present: aviation and health. The basic referral always concerns the individual's medical qualification to fly. All individual's are already selected to fly and most have completed training and are rated. Thus we are not directly concerned with selection though the crew member's return to flying (RTF) is based upon criteria more stringent than those used for initial selection. This is an interesting dilemma. Now that the aircrew member is trained he/she must pass specific standards and tests that were not expected of him/her before selection. Some (mostly rated officers) would say, eliminate the aeromedical standards and eliminate the problem. Others(mostly psychologists and other examiners) suggest better application of the standards before selection. That controversy has been in the literature for several decades and is not likley to be solved in this short paper. The issue here however is the psychological examination of individuals with avation and medical questions. Flying (especially today's inventory and missions) requires specialized abilities and skills from physical and psychological domains. Medically, flying may induce unique conditions and changes in both the soma and psyche. Thus, there are several controversial issues in this unique arena of psychological study and service.

Psychological Data Collection

There will always be controversy about how to evaluate psychological phenomena. Aviation/aerospace psychology is no exception. In the space allowed however, I will only present the current approach used at SAM which is the product of years of study and continues to change. Like other well-educated, achievement orinted and bright , healthy groups (Perry, 1971), aircrew members monitor and screen their self-disclosure (McKinley, Hathaway, and Meehl, 1948). Denial is especially pronounced in our patients because aeromedical consequences are an explicit part of the evaluation. Thus, how does one get accurate information on which to base a fly/no fly decision: interview alone or interview with psychometrics? We practise the latter. Indeed, at least one interview will be based upon the

hypotheses generated from the test data. It would be much more economical to eliminate the testing; and the patients would be delighted to avoid the ten or more hours required for test administration. Matarazzo (1978) argues that the interview lacks sufficient reliability particularly in populations already screened for high achievement. Such selection obviates major classes of psychopathology. Thus, it's not difficult to eliminate major depression, schizophrenia or mental retardation from the ranks, our selection criteria almost accomplish this. Just as worrysome in aviation however, are subtle personality features such as decision making, speed of information processing, stress resiliance, emotional stability, habit control etc. that are hard if not difficult to ellicit information about by talking to a verbal, intelligent, socially skilled person who is heavily invested in keeping this information quiet. Other agencies dealing with equivalent populations such as NASA have elimnated testing. Since the Appolo Project, NASA has relied exclusively on the interview for selection (Jones, & Annes, 1983). Hartman and McNee (1977) suggested that testing works better for medical evaluation than for selection. Our data oriented approach has evolved out of the necessity to make accurate decisions, guard against false-negative bias and out of the belief that, to a point, additional data points add to the accuracy of decisions. The next question however is what psychometrics to add to the interview.

Psychological Tests

Having accepted the importance of testing, the next controversy is whether to use self-report/norm-based tests alone or with projective tests. Both are criticized and both have strengths and liabilities. Both are used at SAM in the attempt to minimize weaknesses and maximize strengths by using a balance of both techniques followed by hypothesis testing in the interview and debriefing. The most salient issues in clincal evaluation are the interpretation problems of projectives and the sensitivity of self-report instruments to faking. These criticisms while valid suggest relying on the balance of both types of instruments. Interpretation is required by both. Self-report tests leave interpretation to the subject while projectives are interpreted by the examiner. Kelly (1958) may state the case best: "When the subject is asked to guess what the exaaminer is thinking, we call it an objective test; when th eexaminer tries to guess what the subject is thinking, we call it a projective device" (p.332). Reliability is tradtionally best with self-report (Wiggins, 1973) but validity can be influenced by test-taking set (Wiggins,1966). Reliability of projective tests is, by psychometric standards low and they are difficult to validate (Exner,1976), but they provide a qualitatively different kind of information about personality that can't be had from a number: modal vs. molecular information. Thus, the clinical utility of both self-report and projective instruments is important to psychological evaluation of aviatiors. However, both kinds of tests are useless unless an individual's results can be compared with the results of similar people.

Norms

Clinical judgment remains something of an art form more or less reliant on science. Norms are part of the science that are particularly important in clinical aviation psychology. Among other things, as a group, aviators are more intelligent, better educated, and more extroverted than groups used to norm tests, including college students (Fine & Hartman, 1968). Thus, their results should be compared with norms more representative of the group. Whether these differences are necessary for flying duty much less psychologically healthy is a question for another paper. Since the aircrew norms tend to be higher than standard norms the subtleties alluded to above could not be evaluated using only general population norms. For instance if an aircrew member's measured intelligence is less than 110 he/she is likely to have trouble keeping up with the cockpit work load, thus such a score would suggest average or slightly below average position in the distribution of aircrew vs. above average ranking in the general population. We continue to norm and update norms on the instruments we use. We have established norms for MMPI, Halstead-Reitan, and WAIS. But the validity of our standard instruments is already strained by the current operational demands. New technological advances will shortly require up-dated and new valid and reliable psychological evaluation techniques.

Operational Requirements

The current challenges to accurate psychological evaluation come from norms and validity. In neuropsychological cases for instance we are asked what does the literature say about a score of 48 on the categories test in relation to cognitive skills necessary to fly the F-16; and how would the answer change if we asked the same question about KC-135 requisite skills? Or further add to the question the history of very mild closed head injury in the absence of neurological findings and with the same categories score: which is borderline for clinical norms and more than two standard deviations above the mean for aircrew. I recommended grounding and reevaluation but the question I faced from my flight surgeon friends was what does the score really mean? Show me the validity trials.

Another current dilemma involves the evaluation of women and racial minorities. We have seen about 5 women in my 3 years at the school, and their results deviated from aircrew norms. Others (   ) have documented a variety of physical and emotional differences that may relate to aircrew duties and suitability. Our numbers are small but are suggestive of a variety of personality differences that may or may not be relevant to flying. All but one of our group were returned to flying based more on clinical judgment than on test norms etc.

Similar to the first question is the measurement of neuropsychological skills for empirical testing of prophylactic

drugs. Do the tests show decrement? If not is that due to absence of effect or lack of sensitivity. Further, if decrement is demonstrated what kind and amount of decrement is acceptable?

We also are faced with the question of single seat, high performance plane vs. crewed, heavy systems and the match or mismatch of man and machine. This is issue is not only a selection problem because we have seen a few pilots whose complaints and symptoms were in part related to the plane his was flying and his difficulty keeping up with it. Our data must be stretched in the best clinical tradition to come to some of these impressions.

All of these and many up-coming questions require empirically based answers and cloud the supposed line between experimental, clinical and human factors psychology.

Conclusion

I have presented an overview of current methods used for psychological evaluation of flyers. I have also shown some of the current controversies surrounding our approach. Further I have discussed a few of the operational and mission-related questions that psychologists are and will be expected to confront. I believe the field of aviation/aerospace psychology is wide open and requires the diverse resources of the science and art of psychology working cooperatively in order to meet the challenges of man in the air and space.

References

Exner, J.E.Projective techniques. In I.B. Weiner (Ed.), Clinical Methods in Psychology. New York: Wiley, 1976.

Fine, P.M., & Hartman, B.O. Psychiatric strenghts and weaknesses of typical Air Force pilots. SAM TR 68-121, USAF School of Aerospace Medicine, AMD, Brooks AFB, 1968.

Hartman, B.O., & McNee, R.C. Psychometric charactistics of astronauts. Aeromedical Review 1-77, Brooks AFB, USAF School of Aerospace Medicine, 1977.

Jones, D.R., & Annes, C.A. The evolution and present status of mental health standards for selection of USAF candidates for space missions. Aviation, Space, and Environmental Medicine, 1983, Aug, 730-734.

Kelly, G.A. The theory and technique of assessment. Annual Review of Psychology, 1958, 323-352.

Matarazzo, J.D. The interview: Its reliability and validity in psychiatric diagnosis. In B.B. Wolman (Ed.), Clinical Diagnosis of Mental Disorders: A Handbook. New York: Plenum Press, 1978

McKinley, J.C., Hathaway, S.R., Meehl, P.E. The MMPI: VI, the k

scale. Journal of Consulting Psychology, 1948, 12, 20-31.

Musgrave, S. Personal reflections of STS-11. Talk presented at USAF School of Aerospace Medicine, Brooks, AFB, 1983.

Perry, C.J.G. Aerospace psychiatry. In H.W. Randel (Ed.), Aerospace Medicine, Baltimore: Williams & Wilkins, 1971.

Wiggins, J.S. Social desireability estimation and "faking good" well. Educational and Psychological Measurement, 1966, 26, 329-341.

Wiggins, J.S. Personality and Prediction: Principles of Personality Assessment. Reading, Mass.: Addison-Wesley, 1973.

Results

The results of the different KW analyses are provided below:

TABLE 1: Age

	GROUP 1	GROUP 2	GROUP 3	CHI^2	DF	P
JOB	302[1]	272	231	5.07	2	.08
SOC	299	287	225	3.12	2	.21

[1]: Mean of the ranks for this group - high mean represents a high score.

GROUP 1: Soldiers over 21 addressing each other; GROUP 2: Soldiers over 21 addressing soldiers under 21; GROUP 3: Soldiers under 21 addressing each other

TABLE 2: Race

	GROUP 1	GROUP 2	CHI^2	DF	P
JOB	266	290	1.29	1	.25
SOC	307	284	1.08	1	.29

GROUP 1: Minority soldiers addressing another minority; GROUP 2: Minority soldiers addressing caucasions or caucasions addressing each other

TABLE 3: Racial Attitudes/Race

	GROUP 1	GROUP 2	GROUP 3	GROUP 4	CHI^2	DF	P
JOB	381	291	258	132	94.96	3	.0001
SOC	367	290	266	163	64.27	3	.0001

GROUP 1: Race differed but both soldiers saw each other as displaying acceptant attitudes toward soldiers of a different minority or ethnic group and soldiers of the same race whose attitudes paralleled; GROUP 2: Race differed but attitude paralleled and was moderately acceptant; GROUP 3: Race differed but attitude paralleled and was not acceptant; GROUP 4: both race and attitude differed.

TABLE 4: Rank

	GROUP 1	GROUP 2	GROUP 3	CHI^2	DF	P
JOB	214	188	160	10.83	2	.004
SOC	200	192	174	2.45	2	.29

GROUP 1: Soldiers of the same rank addressing one another; GROUP 2: Lower ranking soldiers addressing higher ranking soldiers; GROUP 3: Higher ranking soldiers addressing lower ranking soldiers

TABLE 5: Residence Location

	GROUP 1	GROUP 2	GROUP 3	CHI^2	DF	P
JOB	324	280	177	13.01	2	.001
SOC	285	282	185	9.02	2	.01

GROUP 1: Same residence location (barracks or family housing area); GROUP 2: Different residence location; GROUP 3: Both lived in the civilian community

TABLE 6: Marital Status

	GROUP 1	GROUP 2	CHI^2	DF	P
JOB	325	274	3.96	1	.04
SOC	304	275	1.24	1	.26

GROUP 1: Both married; GROUP 2: Both single or marital status differed.

The Relationship Between Substance Use, Demography and Racial/Ethnic Attitudes Related to Performance and Social Attraction

Ron Smith, Ed.D., CPT, MSC
Division of Neuropsychiatry
Walter Reed Army Institute of Research
Washington, DC 20307

Abstract

Data was collected on 275 soldiers assigned to 75 tank crews. Three types of data were collected: self reported drug and alcohol use; demographics; and opinions about fellow crew member's duty performance, interactions with racial/ethnic minorities, and social or interpersonal attractiveness. Analysis indicated that recreational use of drugs or alcohol, demographic similarity, and positive attitudes toward soldiers of another racial or ethnic group were related to high evaluations of duty performance and social attractiveness.

Introduction

An expectation of small tactical unit leaders is maximizing team (crew, squad, etc.) performance (Marshall, 1947). Cohesion, or the bonding together of soldiers so as to sustain their will and committment to one another and mission accomplishment despite environmental barriers - intra-group factions, turbulance, combat demands, mission stress and so on (ARCOST, 1980), is believed to be one of the more important factors in group dynamics that facilitates team performance and protects team members against psychiatric breakdown in combat (Manning & Ingraham, 1983; Ginzberg, 1959; Stouffer, 1949). Furthermore, anything that divides a team into sub-groups or factions undermines or impedes cohesion (Schacter, 1954). The research presented herein examines the relationship between similarity in demography, drug and alcohol use, attitudes toward minorities and opinions soldiers hold of one another concerning job performance and social attraction.

Methodology

Soldiers, assigned to 75 tank crews (totaling 275 soldiers) and stationed in West Germany, were interviewed. The interview consisted of 32 cohesion questions (each with a 5-point Likert scale) querying one soldier about each member of his crew. The questions addressed perceived job performance, social affiliation and racial attitudes. Demographic information as well as self-reported drug and alcohol use (via responses to 5 questions querying frequency of alcohol and drug use in different situations - socially, on duty, during breaks, etc. - each with a 5-point Likert scale) was also collected. The cohesion questions were factor analyzed and two major factors (one labeled job performance (JOB) and the second labeled social attractivness (SOC)) were produced, from which scale scores were derived. That the two primary factors yielded by the factor analysis were labeled duty performance and social attractiveness is consistent with the literature on group dynamics among co-workers (Bennis, 1973). A Kruskall-Wallis (KW) test was performed on the scale scores to see if differences were discernable between groups of subjects who possessed the same or different characteristic (i.e., same age or different age, etc.). The variables used in the KW analysis were age, race and racial attitudes, residence location, rank, marital status, and drug or alcohol use. A detailed presentation of the research design and data analysis is available in Smith (1982a, 1982b, 1983).

TABLE 7: Drug and Alcohol Use

	GROUP							
	1	2	3	4	5	CHI^2	DF	P
JOB*	331	288	259	249	192	20.52	4	.0004
SOC**	326	319	295	280	239	16.42	4	.002

*GROUP 1: Alcohol only and infrequent drug users addressing each other; GROUP 2: Non-users (drugs or alcohol) addressing each other; GROUP 3: Drug users addressing non-users; GROUP 4: Frequent drug users addressing each other; GROUP 5: Non-users addressing drug users.

**GROUP 1: Frequent drug users addressing each other; GROUP 2: Alcohol only and infrequent drug users addressing each other; GROUP 3: Non-users addressing drug users; GROUP 4: Drug users addressing non-users; GROUP 5: Non-users addressing each other.

Discussion

When crew members were asked to rate their fellow crew members along a continuum, differing levels of bonding were found between members of any given crew. Across analyses, when two soldiers shared a trait in common they tended to rate one another high on both factor analytic scales. Likewise, where both soldiers used drugs or alcohol, recreationally but not in excess, they rated one another high on both scales. When both soldiers displayed acceptant attitudes toward minorities or ethnic groups, the ratings were high on both scales whether their race was the same or different. Consequently, crews do appear to be internally divided into factions by variables not easily altered by the soldier or his commander.

Soldiers bond with certain soldiers and not with others. For example, soldiers who live close to one another tend to bond whereas those living far apart do not. Initially, soldiers may bond along more superficial lines defined by common demography. Deeper levels of bonding may be more dependent upon qualitative factors such as attitude similarity. Altering the impact of demography on choices crew members make and the opinions they hold about one another will not be easy. Given the barriers that demography and personal differences provide, only when crew members expect to work together for long periods will they see one another at a deeper level than that afforded by demography.

Fostering intra-unit cohesion will perhaps prove more difficult than fostering inter-racial harmony, for structuring the environment to provide common objectives necessitating a mutually cooperative effort as well as varieties of commonly shared and interdependent experiences is not easily regulated by policy at the institutional level. Yet, minimizing the influence the variables addressed by this research introjects into the crew climate is essential if cohesion is to occur let alone be maintained. Changes at the institutional level are required if the situation at hand is to be effectively alleviated. Stability in crew assignment would be the most desireable solution and probably the most difficult to achieve because of promotions, differences in tour lengths for married and single soldiers, and so on that the unit level manager has little or no influence over. Even so, short term approaches are available as counter-measures to barriers to cohesion until long term changes can be instituted.

Part of this quest is to minimize turbulance in military teams and maximize commonality among soldiers comprising those teams. Turbulance within crews helps to maintain the tendency for soldiers to make preferential choices among their team members resulting in crew division, not unification. Stabilizing unit assignment without stabilizing crew assignment will do little to facilitate cohesion within small tactical military teams.

Replacements will always be a way of life in military units. Even so, making teams more resistent to the disruption caused by the loss and gain of personnel is possible. Therefore, creating a team structure that affords rapid incorporation of new personnel

while retaining as many personnel within the crew as possible is highly desired. Instituting personnel assignment methods capitolizing on naturally ocurring behaviorial tendencies among soldiers could do much to minimize the effects of turbulance and facilitate cohesion at the same time. The current methods result in these same tendencies functioning as barriers to cohesion. The use of common denominators (such as marital status and housing location) having multiple relationships with other commonalities among soldiers (age, rank, after hours socializing, and so on) could be one basis for developing a crew assignment method. Consequently, an examination of the policies and procedures that drive crew assignment and change as well as the personnel tasking requirements placed on units is needed.

Secondly, assigning team members to the same barracks room and assigning members of the same unit to sections of the family housing areas could facilitate cohesion both at the team and unit level. The construction of housing areas that co-locate unit members (instead of separating unit members into barracks dwellers and family housing residents and those who live on the economy) could also facilitate cohesion while minimizing the effects of turbulance. Obviously, building or designing new housing areas is not within the control of the local commander. Yet assigning crew members to the same barracks room is. Housing proximity also has a direct relationship to inter-racial attitude change. The scientific literature indicates that positive inter-racial attitude change is greatest when people of different races both work and live together, less change occurs when they co-reside but do not work together and even less change occurs when they work together but do not co-reside (Deutsch, 1951; Winer, 1952).

Thirdly, leaders must be sensitized to ways of intuitively judging cohesion in their units, trained to build and maintain cohesion, and be given the latitude as well as resources to practice what they are tought and to make alterations consistent with what they see. First line supervisors of small tactical teams are junior leaders (sargents and lieutenants). Training these junior leaders in group process techniques that lend themselves to ready incorportation into every day supervision and by providing them with resources to use as rewards for those they supervise, would also facilitate cohesion at the platoon and crew level. Without these tools, these leaders are actually no different than those they lead.

The squad leader/crew chief or platoon sergeant or leader is the first line supervisor for small tactical teams. Yet, the training given them regarding building and maintaining cohesion is minimal. The day-to-day latitude they have, in their units, to dispense or withold rewards strong enough to shape or maintain the behavior of their subordinates is minimal, as well. The absence of these skills or resources plus crew turbulance and crew members that tend to form alliances within the crew but not as a whole crew makes the the small tactical team leader's job an ominous one.

More senior and experienced leaders in a unit should be available to junior leaders, as a resource. Often, junior leaders do not know what to look for, among their subordinates, as indicators of a lack of cohesion. Secondly, these junior leaders are fearful that asking a senior NCO or Officer for advice will result in their being seen as incompetent. Thirdly, junior leaders must be open to suggestions based in experience, whether the advisor is senior or junior in rank.

Furthermore, incorporating practical approaches to day-to-day supervision relevant to cohesion into all phases of the training given to junior leaders - PLC, NCOES, ROTC, OCS, - is needed. Classroom training alone will not suffice. Once exposed to concepts presented during training, the environment where the leader functions must afford the opportunity to apply this learning and to polish his skills to the point where they become rote. Therefore, an examination of both areas - training and practice - is required and alterations made as needed. Neither area lends itself to a quick fix approach. Both must be methodically examined and the appropriate changes made from an organizational viewpoint, integrated and implemented at the unit and instructional level.

Summary

Small military tactical teams are composed of soldiers who are highly diverse. Such diversity potentiates barriers to cohesion's development and maintenance. The quest is to find countermeasures for naturally occurring barriers that demographic differences and personal traits as well as attitudes or habits introduce into the team's climate and/or institute a team structure resistent to them.

Similarity in demography, personal habits and attidudes is directly related to the degree of bonding between any two soldiers. Yet, combat teams are formed solely on the basis of rank and MOS. Given the barriers that demography provides, only when crews expect to work together for long periods will team members see one another at a deeper level than that afforded by demography. Turbulance coupled with interpersonal incompatability within crews helps to maintain the tendency for soldiers to make preferential choices among their team members resulting in crew division, not unification. Stabilizing unit assignment without stabilizing crew assignments will do little to facilitate cohesion among team members.

By examining the demography of crews short personnel and assigning new crew members accordingly, even by tour length, more intra-crew compatability and less turbulance would occur. Tour length is closely related to marital status and rank, which correlate highly with age and residence location among junior enlisted personnel. By so doing, a few squads or crews would be clearly mismatched. Yet, under current methods, all crews are so characterised.

References

ARCOST. Improving stability and enhancing cohesion in the US Army. US Army Deputy Chief of Staff for Personnel. 1980.

Bennis, W. Interpersonal dynamics. Homewood, IL: Dorsey. 1973.

Deutsch, M. Interracial housing. Minneapolis: University of Minnesota Press. 1951.

Ginzberg, E. The ineffective soldier. New York: Columbia University Press. 1959.

Manning, F., Ingraham, L. An investigation into the value of unit cohesion in peacetime. Washington, DC: Division of Neuropsychiatry, Walter Reed Army Institute of Research. 1983.

Marshall, S. Men against fire. New York: William Morrow. 1947.

Schacter, S. An experimental study of cohesiveness and productivity. Human Relations, 7: 403-439, 1954.

Smith, R. The impact of drug abuse on tank crew cohesion. Alexandria, VA: Defense Documentation Center. 1982a.

Smith, R. Cohesion in military work teams. Division of Neuropsychiatry, Walter Reed Army Institute of Research, Washington, D.C. 1982b.

Smith, R. The buddy relationship. Division of Neuropsychiatry, Walter Reed Army Institute of Research, Washington, D.C. 1983.

Stouffer, S. Studies in the social psychology in World War II. Princeton; Princeton University Press. 1950.

Winer, D. Residential proximity and intergroup relations in public housing projects. Journal of Social Issues, 8(1): 45-69, 1952

Disclaimer

AD-P003 288

Attitudes Toward Making A Transfer: A Predictive Model

James B. Shaw
Cynthia D. Fisher
Richard W. Woodman

Texas A & M University

Abstract

This report represents a study of 143 USAF personnel who had been notified that they would be making a permanent change of station (PCS). Respondents came from 7 USAF bases in 3 states. Data were collected on attitude toward the upcoming PCS as well as on 8 variables that were thought to predict PCS attitude. They were: relative similarity of present and future assignments, frequency of past transfers, extent to which the new assignment was seen as a career advance, attractiveness of the new assignment, attractiveness of the old assignment, extent of successful adjustment to past transfers, notice time given prior to transfer, and ratio of preferred PCSs to total career PCSs. Five of the 8 variables correlated significantly and in the predicted direction with PCS attitude. Regression analyses yielded a multiple R of .68.

Introduction

This report presents the results of a study of 143 USAF NCO personnel who had been notified that they would make a PCS. Data were collected on attitude toward the PCS and 8 variables thought to predict that attitude.

Similarity of the move. Based upon the work of Louis (1980) and Brett and Werbel (1980), one hypothesis tested was that as the differences between the new situation and the present assignment increased, the individual's fear of the unknown would increase, and reluctance to make the move would increase. However, individuals who are located in an undesirable geographic or job situation may view a move to a very different situation very positively. Thus the similarity of the move may interact with the individual's attitude toward the present situation.

Transfer frequency. Edstrom and Galbraith (1977) suggest that frequent transfers increase an individual's commitment to the organization. Brett and Werbel (1980) found a positive relationship between frequency of moves and job involvement. Fisher et al (1982) suggest that families who choose to accept multiple moves may well be those who are more able to adjust to transfers and/or may have acquired more moving skills over the years. However, the relationship between frequency and attitude may prove to be curvilinear in nature. Seidenberg (1973) suggests that while individuals can cope with several transfers during their life, at some point they may simply "run out of steam" and not be able to handle further moves.

Perceived advancement. Brett and Werbel (1980) and Hammer and Vardi (1981) found that transfers were more readily accepted if they were viewed as enhancing the individual's career. We tested the hypothesis that as an individual's perception of the transfer as a significant career step increases, attitude toward making the transfer becomes more positive.

Attractiveness of the new assignment. Pinder (1980) and Seidenberg (1973) found that the attractiveness of the geographical location of the transfer was positively related to a family's likelihood of accepting the transfer. Baker (1976), Burke (1974), Levenson and Hollman (1980), and Pinder (1980) found that the attractiveness of the new cultural environment and general life style related postively to family attitude toward the transfer.

Attractiveness of the present assignment. Burke (1972,1974) found that a family's satisfaction with a transfer location was determined by their feelings toward their old community. We hypothesized that as satisfaction with the old community and/or job increased, the attitude toward the transfer would become less positive.

Past transfer adjustment success. Brett and Werbel (1980) found that the wives of individuals making a transfer were more willing to move when their last transfer had been successful. We hypothesized, then, that to the extent that individuals and their families have succeeded in adjusting to previous transfers, their attitudes toward other transfers will become more positive.

Notice given before the transfer. Burke (1974) found that notice time given to employees to make a move was positively related to satisfaction with the assignment. Pinder (1979) found that notice time predicted an employee's attitude toward company transfer policies. Fisher et al. (1982) suggest that notice time affects the extent to which employees can make an "informed decision" and may influence subsequent satisfaction.

Transfer history. In the USAF, the choice an individual has in transfer decisions is somewhat limited. Some individuals have a history of "desirable" transfers while others have often been transferred to "undesirable assignments." From equity theory (Adams, 1963), one reasonable hypothesis is that the ratio of desirable transfers to total career transfers is positively related to transfer attitude.

Method

A total of 143 male USAF NCOs were interviewed using a survey instrument which included both fixed-response and open-ended items. Data were collected in April and May of 1983.

Sample

The NCOs who participated in the study represented 81 job classifications and were stationed at one of 7 USAF bases in the South and Southwest. The individuals had received notice that they would be making a PCS. Subjects were selected based upon their past transfer history so that a broad range of individuals, some of whom had been transferred frequently and others who had been transferred infrequently, would be included in the sample. Our sample selection process attempted to minimize the relationship between transfer frequency and tenure in the USAF by including some longer term personnel who had been moved infrequently.

Survey Instrument

The survey instrument consisted of three sections with a total of 105 items. Sections 1 and 3 were completed individually by the respondent, while the second and largest part of the survey was completed using an interview format.

Dependent Variable: Attitude Toward Transfer

Two items measured attitude toward the PCS. One was a 7 pt. (1=very reluctant, 7=very eager) fixed-response item, while the second item consisted of five 7 pt. semantic differential scales. Attitude toward the PCS was measured by summing all five semantic differentials and the 7 pt. fixed-response item.

Independent Variables

Similarity of the move. Five indexes of move similarity were obtained. Similarity of the new assignment to the present in terms of climate, city size, culture, and job were assessed using a 5 pt. "not similar at all" to "the same" scale. Another item dealt with the degree to which the new job had more or less responsibility than the present job. Having more or less responsibilty were both regarded as decreasing the similarity of the move. An overall similarity score was derived by summing the responses for all five aspects of the transfer (alpha = .34). In addition, an "interaction" score was derived by multiplying the overall similarity score by the overall match of the present location to an ideal (see below).

Transfer frequency. Six indexes of transfer frequency were obtained: total number of PCS moves, average number of months per move, number of months since the last PCS, number of career short tours, total days TDY in the last 12 months, and the number of TDY assignments over the last 12 months.

Perceived advancement. Perceived advancement was measured using a 7 pt. Likert style item.

Attractiveness of the new assignment. One item measured the degree to which the climate, city size, culture, and geographical location of the new assignment matched the individual's "ideal" for these four characteristics. Job assignment "match" was also measured. An overall match score was derived by summing the scores on the two items (alpha = .78). Another item was used to measure the expected satisfaction associated with the new assignment. Expected satisfaction with the job, co-workers, base, housing, and community were measured. An overall satisfaction score was derived by summing across the five facets (alpha = .74).

Attractiveness of present assignment. Items almost identical to those used to measure expected satisfaction were used to measure satisfaction with the present assignment and similar items were used as above to measure "match" of the present assignment to ideal. Overall match (alpha = .73) and satisfaction (alpha = .56) scores were obtained as with attractiveness of the new assignment.

Past transfer adjustment. Rated family adjustment difficulty was measured using a 7 pt. "very difficult" to "very easy" scale. A rated personal adjustment score was obtained using the same scale. A personal weeks to adjust score was computed by averaging the responses to three items which asked how long it had taken the individual to adjust (in weeks) to the community, technical aspects of their job, and co-workers.

Notice given before transfer. The amount of notice time in months given prior to transfer was measured using two items: "when were you notified" and "when will you depart."

Transfer history. Respondents were asked to list each transfer they had made during their career and whether they had wanted to make each transfer. This information was used to compute the ratio of preferred moves to total career moves.

Results

Eight hypotheses were tested using correlational analysis procedures. Regression procedures were used to test the capabilities of the 8 major independent variables to predict PCS attitude. Data were availible for less than the total sample of 143 due to non-responses on some items.

Test Of Major Hypotheses

The results of the correlational analyses which provided the major test of our hypotheses are presented below. Where several indexes were summed to form an overall score, only the overall score is presented unless results related to a specific index are dramatically different from the overall score.

Table 1: Correlations Among Major Independent Variables And PCS Attitude

Independent Variable	r
1. Similarity of Present To Next Move:	
Overall Simlarity	-.04
Similarity X Present Assignment Match	-.04
2. Transfer Frequency:	
Total # Of Career Transfers	.01
Average # Of Months Per PCS	-.06
Time Since Last PCS	-.21 *
# Of Career Short Tours	-.01
Days TDY In Last 12 Months	.15
# Of TDY Assignments In Last 12 Months	.11
3. Perceived Advancement	.39 **
4. Attractiveness Of New Assignment:	
Overall Match Score	.53 ***
Overall Expected Satisfaction Score	.50 ***
5. Attractiveness Of Present Assignment:	
Overall Match Score	-.23 **
Overall Satisfaction Score	-.05
6. Ease Of Personal And Family Adjustment In Last PCS:	
Difficulty Of Family Adjustment	.08
Difficulty Of Personal Adjustment	.16
# Of Weeks Needed For Personal Adjustment	-.32 ***
7. # Of Months Notice Time Prior To Transfer	-.11
8. Ratio Of Preferred/Total Career PCSs	.26 **

* $p < .05$ ** $p < .01$ *** $p < .001$

Similarity of the move. Hypothesis one stated that there would be a positive correlation between similarity and PCS attitude. No support for this hypothesis was found. Also, no significant correlation was found between an interaction score and attittude.

Transfer frequency. Hypothesis two predicted a positive relationship between transfer frequency and PCS attitude. Only one of the six indexes of transfer frequency significantly correlated with PCS attitude. A significant negative correlation was found between "time since last PCS" and PCS attitude. This finding is consistent with hypothesis two, since the negative correlation indicates that as the time since last PCS increases (infrequent transfer), the attitude toward the PCS becomes less positive.

Perceived advancement. A very strong positive correlation was found between perceived advancement and PCS attitude.

Attractiveness of the new assignment. Overall "match to ideal" was significantly and positively correlated with PCS attitude, as was expected satisfaction with the new assignment.

Attractiveness of the present assignment. The extent to which the present assignment matched the individual's ideal was correlated significantly and in the predicted negative direction with PCS attitude. No significant correlations were found between satisfaction with any of the five aspects of the present situation and PCS attitude.

Past transfer success. PCS attitude correlated significantly with the number of weeks needed for personal adjustment to the community, job, and co-workers in the last PCS. The longer it took the individual to successfully adjust to their last PCS, the less positive their attitude toward the upcoming move.

Notice given before transfer. The correlation between months notice and PCS attitude was not significant, and was, in fact, in the opposite direction from that predicted.

Transfer history. Hypothesis eight was supported in that the correlation between PCS attitude and the ratio of preferred assignments to total career assignments was statistically significant and in the predicted direction.

Predictive model of PCS attitude. Stepwise regression procedures were used to examine the predictive capacities of the independent variables. Those variables which showed a significant ($p < .05$) correlation with PCS attitude were used in the regression analysis. Where composite scores were availible, these were used rather than the individual indexes that comprised that overall score. The results of the regression analysis are presented below. An overall multiple R of .68 was achieved with five variables accounting for 46% of the variance in PCS attitude.

Table 2: Results Of Regression Analysis

Variable	R	R Square	Beta	F
Overall Match Of Next Assignment To Ideal	.51	.26	.396	16.87 **
Overall Match Of Present Assignment To Ideal	.58	.34	-.316	15.99 **
Ratio Of Preferred/Total Career Assignments	.65	.42	.222	7.26 **
Expected Overall Satisfaction	.67	.45	.212	4.80 **
# Of Weeks Needed For Personal Adjustment	.68	.46	-.137	2.96 *

* $p < .05$ ** $p < .01$ N = 97

Discussion

A finding of no relationship between similarity of the new assignment to the old and PCS attitude is inconsistent with the work of Brett and Werbel (1980) and Louis (1980). However, these authors suggested that the similarity of the transfer assignment would affect transfer adjustment, not necessarily attitude toward the transfer. To examine this issue, data concerning the similarity of the previous job assignment and present job assignment were collected. Data were also collected concerning how many weeks it had taken the individual to "get up to speed" and to adjust to the technical aspects" of their present job assignment. We also had a measure of how long the individual had been in their present assignment. Our sample was then divided into two groups: (1) those who had rated their previous job as being "not similar at all" to their present job assignment, and (2) those who had rated their previous jobs as being "similar" to "the same" as their present job. A "percent adjustment time" score was computed by dividing the number of weeks the individual had needed to both "get up to speed" and "adjust to the technical aspects" of their job by the length of their present assignment in weeks. Data indicated that the Low Similarity individuals spent 22% of their time getting up to speed in their present job and 16% of their time getting adjusted to the technical aspects of their job. High similarity individuals spent only 10% and 6% respectively. No differences were found between the two groups in terms of the length of their present assignment. This data supports earlier research, and indicates that a significant "productivity loss" occurs in transfers associated with significant changes in job duties.

Data concerning hypothesis two and others seem to suggest that the history of transfer frequency over the entire career of the individual has little effect on attitude toward a specific upcoming PCS. The process involves more of a direct comparison of what has occurred in the present assignment and what is likely to occur in the next. A major determinant of PCS attitude is how long the individual has been in their present assignment. For individuals who have moved frequently in their career, but have been at their present assignment for some considerable number of years, PCS attitude is more related to the disruption of their present assignment than by the skills and perspectives they may have gained from earlier and more frequent moves. No significant correlations were found between PCS attitude and either measure of TDY. This may be because of the considerable restriction of range found in the TDY measures.

The positive relationship of PCS attitude to perceived advancement and attractiveness of the new assignment, and the negative correlation between PCS attitude and attractiveness of the present assignment are consistent with earlier work. The finding that the degree to which the present assignment "matched" an ideal

assignment was correlated with PCS attitude, but satisfaction with the present assignment did not, was very interesting. Individuals in the sample distinguished between being satisfied with an assignment vs. viewing the assignment as "ideal." Individuals might view any number of assignments as potentially satisfactory, but only a very few as ideal. Thus, transfer from an ideal assignment would have a greater negative effect on PCS attitude that transfer from an assignment that is simply satisfactory.

PCS attitude was correlated with number of weeks to personally adjust to the new community, job, and interpersonal aspects of the job, but not with ratings of family adjustment or personal adjustment difficulty. Individuals may have been able to more accurately portray their adjustment process in terms of "how long it took them" rather than simply "how difficult it was." Also, in the measure of PCS attitude, it was assumed that the individual would "factor in" the attitudes of his family in replying to the item. This may not have occurred.

Finally, related to hypothesis eight, a postive relationship was found between the number of preferred/total career PCSs and PCS attitude. The magnitude is modest ($r = -.26$) and suggests that a more complex relationship between preference and attitude exists. Individuals who have a high preferred/total ratio might view transfer to an undesirable location very negatively, since they are use to getting good assignments. On the other hand, according to equity theory (Adams, 1963), a high ratio of preferred/total PCSs would lead to a greater acceptance of the undesirable PCS due to internally and/or externally derived notions of fairness. Our study indicates that these equity processes might be very important in understanding PCS attitude, but our data did not allow us to examine these issues adequately.

The results of the regression analysis were very encouraging. With five variables in the equation, 46% of the variance in PCS attitude was predictable. Because of the small sample size, the results certainly need to be replicated. However, the results do suggest that carefully designed survey instruments which measure key factors such as those included in the present study, could prove very helpful in identifying those individuals whose attitude toward certain PCS moves would be negative and perhaps cause the individual to leave military service rather than accept a transfer.

References

Baker, J.C. Company policies and executive's wives abroad. Industrial Relations, 1976, 15, 343-348.

Brett, J.M. and Werbel, J.D. The effect of job transfer on employees and their families. Report to the Employee Relocation Council, Washington, D.C., 1980.

Burke, R.J. Quality of organizational life: The effects of personnel job transfers. Proceedings of the 32nd Annual Meeting of the Academy of Management, 1972, 32, 242-245.

Burke, R.J. Personnel job transfers: Some data and recommendations. Studies In Personnel Psychology, 1974, 6, 35-46.

Edstrom, A. and Galbraith, J.R. Transfer of managers as a control strategy in multinational organizations. Administrative Sciences Quarterly, 1977, 22, 248-263.

Fisher, C.D., Wilkins, C. and Eulberg, J. Transfer transitions. Report prepared for the Office of Naval Research, Texas A & M University, 1982, Report #5.

Hammer, T.H. and Vardi, Y. Locus of control and career self-management among non-supervisory employees in industrial settings. Journal of Vocational Behavior, 1981, 18, 13-29.

Levenson, M.K. and Hollman, R.W. Personal support services in corporate relocation programs. The Personnel Administrator, September 1980, 45-51.

Louis M.R. Career transitions: Varieties and commonalities. Academy of Management Review, 1980, 5, 329-340.

Pinder, C.C. Corporate transfer policy: Comparative reaction of managers and their spouses. Relations Industrielles, 1979, 33, 654-665.

Pinder, C.C. Mobility and transfer. In H. Meltzer and W. Nord (Eds.) Making Organizations Humane and Productive. New York: Wiley, 1980.

Seidenberg, R. Corporate wives-corporate casualties? New York: AMACOM, 1973.

AD-P003 289

A Path Analytic Study of Psychological Success Among Air Force Officers

Sharon C. Conley

Abstract

This study provided a partial test of Hall's model of psychological success in work organizations. The model proposes that an individual's feeling of task success (psychological success) leads to a "self-perpetuating cycle of success" where the individual is motivated to set goals and perform successfully in his/her work. Previous tests of this model have been conducted in school settings (Hall and Hall, 1976; Hall and Foster, 1977), thereby limiting the ability to generalize the results to work settings.

The sample consisted of 110 Air Force lieutenants who were employed in a variety of Air Force specializations. Path analytic results showed that the data did not achieve a close fit with the theoretical model. By estimating path coefficients from a model that assumed Hall's predictive ordering, but had no ommitted linkages, a model more consistent with the data was developed. The new model depicts satisfaction as an intervening variable between psychological success and motivation instead of between self esteem and job involvement, as specified by Hall.

Introduction

Hall's (1976) model of psychological success in work organizations proposes that feelings of job success lead to an increase in self esteem. Increased self esteem fosters a sense of satisfaction with one's work, which leads to a higher degree of job involvement. The effect of this chain of events is to increase the individual's motivation to set future work-related goals; therefore, the model depicts a self perpetuating cycle.

Partial support for the model has been reported with samples of school children (Hall and Hall, 1976) and college students (Hall and Foster, 1977). However, because neither of these studies were conducted in work settings, the applicability of the findings to occupational experiences is in question. Using a sample of Air Force officers, this research will conduct a partial test of the following causal model developed by Hall: psychological success >> self esteem >> satisfaction >> job involvement >> motivation.

Method

In May of 1982, questionnaires measuring psychological success, self esteem, satisfaction, job involvement and motivation were mailed to 172 Air Force second lieutenants who were employed in different job specialties. One hundred and ten officers responded

(53 males and 57 females), yielding a response rate of 63%. The subjects were originally selected by Air Force Academy researchers for participation in a four year study concerning the integration of women into the Academy. Thus, all of the officers in this sample are graduates of the Air Force Academy.

Self esteem and job involvement were measured with items selected from Rosenberg's Self Esteem Scale and Lodahl and Kejner's Job Involvement Scale, respectively (Rosenberg, 1965; Lodahl and Kejner, 1965). Items used by Hall were adapted to measure psychological success. Hackman and Oldham's Job Diagnostic Survey was utilized to measure job motivation and satisfaction (Hackman and Oldham, 1975).

The internal consistency reliabilities (coefficient alphas) of the scales ranged from .56 (job involvement) to .85 (satisfaction), suggesting that the internal consistency of the scales is generally satisfactory.

Results

Table 1 presents the means and standard deviations for each of the five scales. The means indicate that the average levels of psychological success, self esteem, satisfaction and motivation are high. For job involvement, the average value of 3.8 indicates moderate levels of job involvement. The restricted range on the self esteem scale indicates that the officers are largely homogeneous with respect to self esteem.

Path analysis was utilized to assess the fit of the data to Hall's causal model: (1) Psychological success >> (2) Self esteem >> (3) Satisfaction >> (4) Job involvement >> (5) Motivation. The path coefficients in the model are equal to the zero order correlations because each variable is conceived to be affected by one variable only. Therefore, for the present model:

$r_{12}=p_{21}=\beta_{21}$ (psychological success -> self esteem)

$r_{[illegible]3}$ [illegible] $=\beta_{32}$ (self esteem -> satisfaction)

[illegible] (satisfaction -> job involvement)

[illegible] (job involvement -> motivation)

Figure 1 shows the path coefficients among the variables. All but one of the hypothesized relationships are significant. The non-significant path coefficient, (p32) indicates that self esteem does not have a direct effect on satisfaction. This leads to the

conclusion that this particular path should be deleted from the model. The deletion results in two unrelated models, as shown in Figure 2.

Now the question is, to what extent does the data fit this more parsimonious specification? Kerlinger and Pedhazur (1973) suggest that path coefficients be calculated in an attempt to reproduce the original correlations. This serves as an indication of consistency between the data and the model. According to Kerlinger and Pedhazur's specifications, the following equations represent the zero order correlations for the two models shown in Figure 2:

Model 1: $r_{12}=p_{21}=.39$

Model 2: $r_{34}=p_{43}=.35$

$r_{45}=p_{45}=.36$

$r_{35}=(1/N)x\sum z_3z_5=(1/N)x\sum z_3(p_{54}z_4)=p_{54}r_{43}$

substituting p_{43} for r_{43}:

$r_{35}=p_{54}p_{43}=(.36)x(.35)=.126$

The original r_{35} is .65.

Since the discrepancy between the original and the reproduced model is large (.126 vs .65), it is concluded that the data are inconsistent with the model.

Because of these results, an attempt was made to estimate causal relationships from a model which assumed Hall's predictive ordering, but had no missing linkages. In this model, each variable had as predictors those variables that preceded it in the model. Thus, all of the other variables were predictors of internal motivation.

Path coefficients were estimated for all of the dependent variables. Because path coefficients are equivalent to standardized regression coefficients, regression analysis was used to calculate the path coefficients. Non-significant paths were deleted, and the resulting model is shown in Figure 3.

Once again, the correlation coefficients were reproduced from the new model, and compared to the original correlations. The fit is an improvement from the earlier revision of the model, although it is not extremely good. The discrepancies ranged from a high of .13 between satisfaction and self esteem, to a low of .03 between psychological success and job involvement. Therefore, the model shown in Figure 3 is considerably more consistent with the data than is the model in Figure 2.

Discussion

The predictive relationships that were demonstrated in this study may have been partially due to the personality characteristics of the Air Force officers and to the nature of their work. In general, the officers are in "enriched" jobs that provide them with high levels of significance, variety and feedback (Conley, 1983). In addition, these officers have been characterized as having high needs for personal growth and development and high achievement motivation (De Fleur, 1981; Conley, 1983). People with high growth needs are thought to value the intrinsic rewards that result from performance on challenging tasks (Hackman and Oldham, 1976). Therefore, these results may have been partly due to the high quality of the jobs — and to the values that the officers place on personal growth and achievement.

The results offer partial support for Hall's model. Specifically, the links between self esteem and satisfaction — and between job involvement and motivation — that are specified in Hall's model are not supported. However, psychological success relates to self esteem and satisfaction relates to job involvement as Hall's model suggests. The new model that was developed in this study suggests that satisfaction is an intervening variable between psychological success and motivation, instead of between self esteem and job involvement as specified by Hall.

It should be emphasized that because the new model assumes Hall's predictive ordering, alternative models of the causal relationships can not be ruled out.

References

Conley, S. The relationships among expectations, perceived job characteristics, and work outcomes. Unpublished doctoral dissertation, University of Michigan, 1983.

De Fleur, L.B. Report submitted to the Air Force Office of Scientific Research, 1981.

Hackman, J.R., and Oldham, G.R. Development of the job diagnostic survey. Journal of Applied Psychology, 1975, 60, 159-170.

Hackman, J.R., and Oldham, G.R. Motivation through the design of work: test of a theory. Organizational Behavior and Human Performance, 1976, 16, 250-279.

Hall, D.T. Careers in organizations. Pacific Palisades, Ca.: Goodyear, 1976.

Hall, D.T., and Foster, L.W. A psychological success cycle and goal setting: goals, performance and attitudes. Academy of Management Journal, 1977, 20, 282-290.

Hall, D.T., and Hall, F.S. The relationship between goals, performance, success, self image and involvement under different organization climates. Journal of Vocational Behavior, 1976, 9, 267-278.

Kerlinger, F.N., and Pedhazur, E.. Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston, 1973.

Lodahl, T.M., and Kejner, M. The definition and measurement of job involvement. Journal of Applied Psychology, 1965, 49, 24-33.

Rosenberg, M. Society and the adolescent self image. Princeton, N.J.: Princeton University Press, 1965.

Table 1

MEANS AND STANDARD DEVIATIONS FOR PSYCHOLOGICAL SUCCESS, JOB INVOLVEMENT, SELF ESTEEM, SATISFACTION AND MOTIVATION

	Mean	SD	Range
Psychological success	5.6	1.06	1.0-7.0
Self esteem	4.5	.53	2.7-5.0
Satisfaction	5.0	1.53	1.0-7.0
Job Involvement	3.8	1.35	1.0-7.0
Motivation	5.9	1.05	1.0-7.0

[a]N=110

[b]7-point response scales were used for all measures except self esteem, where a 5-point response scale was used.

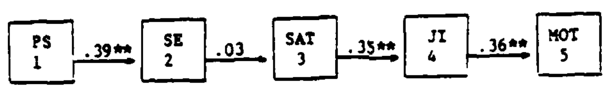


Figure 1. A path diagram representing the causal relationships in Hall's model of psychological success.

**p <.01

Figure 2. Revised path diagram of psychological success.

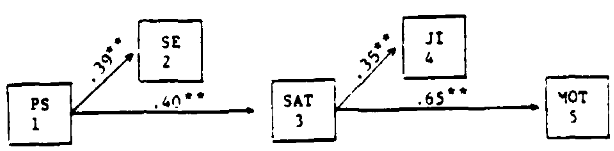


Figure 3. Significant path coefficients from a model with no ommitted linkages.

**p <.01

SYMPOSIUM INVITED ADDRESS

VISUAL DISPLAY TECHNOLOGY-A DEVELOPING CHALLENGE FOR THE BEHAVIORAL SCIENTIST

Dr. Harry L. Snyder
(Virginia Polytechnic Institute and State University)

Harry L. Snyder is the R. H. Bogle Professor of Industrial Engineering and Operations Research at the Virgina Polytechnic Institute and State University. He has taught there since 1970, and between 1975 and 1979 he served as head of the Department of Industrial Engineering and Operations Research. He received his undergraduate degree with high honors in psychology from Brown University in 1958, and completed his M.A. (1960) and PhD (1961) in experimental psychology from the Johns Hopkins University.

Professor Snyder's areas of research and professional interest include human visual performance, display design and evaluation, statistics and experimental design, and human factors in product liability. He has been the Principal Investigator on numerous projects, including a wide range of research conducted for agencies within the Department of Defense. His work has been widely published in The Journal of Experimental Psychology, Human Factors, Ergonomics, and many other professional journals; he has also presented his work at many national and international symposia and professional meetings. His professional affiliations include membership and fellowship in the Human Factors Society, the Society of Engineering Psychologists, the Optical Society of America, as well as membership in the Ergonomics Society, the Society of Photo-Optical Instrumentation Engineers, and the American Institute of Industrial Engineers. In addition, Dr. Snyder is a private pilot, instrument rated.

His invited address at the Ninth Biennial Psychology in the DOD Symposium on 19 April 1984 is titled "Visual Display Technology - A Development Challenge for the Behavioral Scientist." A summary of his address appears on the following pages.

Visual Display Technology:
A Developing Challenge for the Behavioral Scientist

Harry L. Snyder
Virginia Polytechnic Institute and State University

Abstract

Visual display technology is developing at a pace far exceeding that of the required behavioral data for representative task requirements. Relations between current and projected display capabilities are compared with pertinent human visual performance data and limitations.

Introduction

One of the frequently cited frustrations of the human factors engineer or engineering psychologist is that the design process has progressed too far and too fast before he or she is called to participate. This issue has been raised in the development of numerous military systems, has been apparent in far too many industrial and civilian systems, and raises its unfortunate head in nearly all technology development areas in which the technology has a critical relationship to the human operator. Visual display technology is no exception.

Visual displays are being developed in many high technology nations for both civilian and military products. The Department of Defense is a prime force in pushing these developments, particularly in areas dealing with higher resolution, full color, and large size displays. Many of the development efforts are generic and not system specific. In other situations, the intended application is known and the hardware is being pushed to new capabilities without analysis of the requirements and limitations of the human user. Unfortunately, some of these concepts lead to incompatible "solutions," as when color may be a hindrance, too much resolution may yield lower legibility, or too large a display may compromise design of the total user environment. These issues and pertinent data from the fields of engineering psychology and applied visual science are discussed below to put into perspective the asynchrony between human requirements and technology development, and to urge attention to be given toward resolving this problem.

User Requirements and Limitations

Of particular concern in the current fast pace of display technology development are four key parameters of visual displays: size, resolution, uniformity, and color.

Display Size

Many user environments logically and rightfully call for simultaneous viewing of displays by several persons. Examples are command and control facilities, both at sea and on land. To achieve this multiple-user capability, particularly for pictorial status displays, it is meaningful to use physically large displays capable of displaying large amounts of information.

Merely increasing the display size such that the visual angular subtense of information presented on a large display is equal to that of a smaller display,

viewed from a smaller distance, can have some advantages. One of these advantages is that the visual acuity of the eye improves substantially from typical one-person viewing distances of one-half meter to larger distances, on the order of two or three meters (De Palma and Lowry, 1962). Thus, by removing the display to a greater distance, visual performance can be facilitated. On the other hand, larger size displays are more restrictive to flexible workplace design. As we move toward greater usage of very large flat-panel displays, the power, weight, and structural requirements for this display will preclude its relocation in a given facility without considerable effort.

But perhaps more important is the attendant visual effect of off-axis viewing required for multiple viewers. If two persons are view simultaneously a typical small display, they will generally view it from about 45 degrees off axis due to their shoulder widths when seated. This off-axis viewing, even for displays which are advertised to produce "wide-angle viewing," will result in a significantly longer search time for menu items (Snyder and Maddox, 1978). At the present time, there seems to be no limit to the desire of display technologists to push toward even larger displays for multi-viewer environments. For example, the popular AC-driven orange colored plasma panel has found great acceptance in its original size of 8.5 inches square, with 512 by 512 picture elements In more recent years, it has been enlarged to 17 inches square (1024 by 1024 pixels), to a 1-meter square version, and most recently to a 2-meter diagonal size version. While the technology improvement is impressive, the resulting use limitations and impact upon user visual task performance are yet to be determined.

Resolution

Since the introduction of commercial television in the 1940s, the resolution of the familiar cathode-ray tube (CRT) has increased at an accelerating rate (Figure 1). Thus, while earlier commercial television systems were capable of displaying about 300 vertical by 400 horizontal pixels, using a 30-Hz frame rate with an interlace to reduce flicker, we have passed through the 1260 vertical by 1600 horizontal pixel resolution in 1970 and the theoretical 2130 by 2840 pixel resolution level in 1980. It is worth noting that we have not seen displays of the 2130 by 2840 resolution variety because the CRT spot size has not been obtained yet, not because the video bandwidth is not available. As a result, manufacturers of these high resolution systems opted for a 60-Hz frame rate, to avoid single pixel flicker, and offerred a resolution of about 1350 by 1800. Most recently, high quality systems have a 60-Hz frame rate with resolution of 1500 by 2000 picture elements. It can be predicted (Figure 1) that by 1990 we will have the resolution capability of 1850 by 2500 pixels, and that by 1995 the resolution of 2000 by 2700 may well be achievable on CRT displays. The real question is not the feasibility but rather the utility of this resolution, a question that relates to the visual system of the observer and not the video bandwidth of the circuitry. For this answer, we need to turn to visual capability data, such as the contrast threshold function (CTF) illustrated in Figure 2.

The CTF reaches a maximum of about 60 cycles/degree at unity modulation under high luminance, but only about 30 cycles/degree at 10% modulation. At maximum resolution, a CRT is unlikely to exceed 10% modulation, so the required resolution of the display need not exceed the equivalent of 30 cycles/degree at the nearest viewing distance. Assuming a single operator at a near distance of 16 inches, the effective display resolution is limited by human vision to the classical rule of thumb of 1 arcminute (equal to 30 cycles/degree) or .0046 inch. To display 2000

FIGURE 1. PROJECTED CRT RESOLUTION

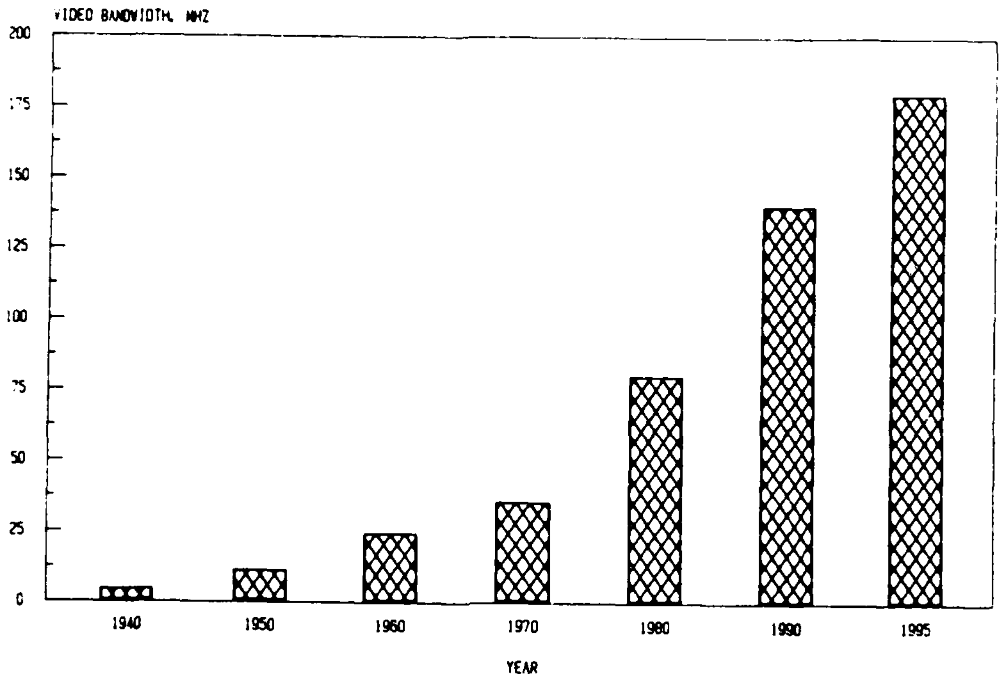
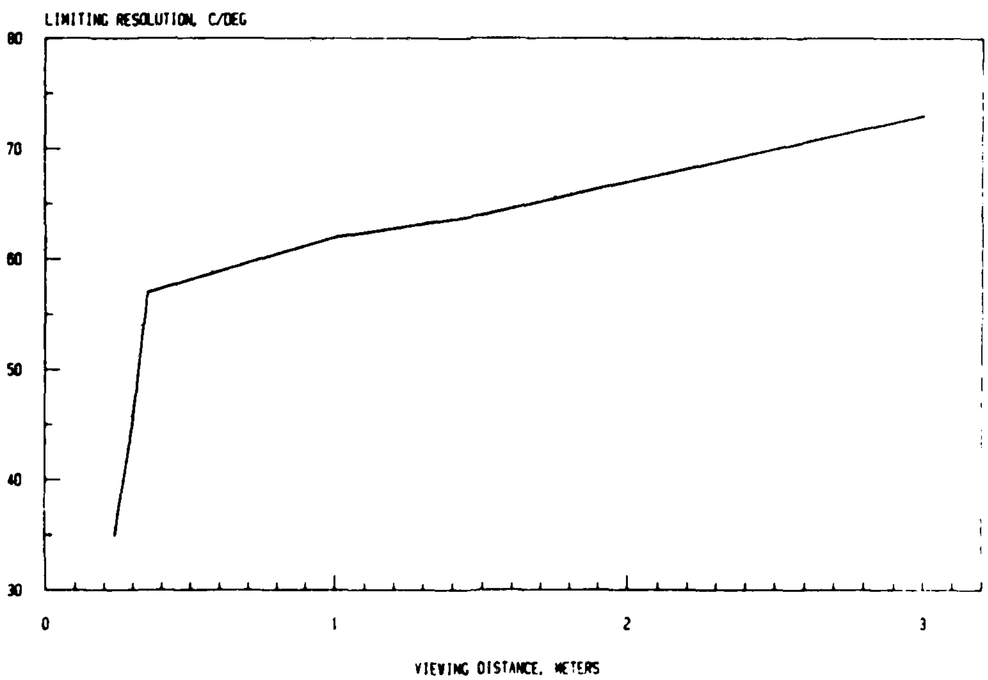


FIGURE 2. VISUAL LIMITING RESOLUTION

EFFECT OF VIEWING DISTANCE



picture elements, each of which is .0046 inch wide, a display needs to be at least 9.3 inches in the horizontal dimension and 7.0 inches in the vertical. At the more typical design viewing distance of 28 inches, the display need not have a resolution better than .008 inch and must therefore be at least 16.3 inches horizontal by 12.2 inches vertical to avoid visual resolution limitations. That is, it must be at least as large as a conventional 21-inch television CRT. How many military aircraft cockpits or other display environments permit this size luxury? Yet, unless the CRT is of that size, the CRT resolution exceeds that usable by the eye and therefore is more costly than needed. Certainly, except for larger displays or those viewed at closer distances, there seems to be little need for further resolution development of CRTs.

The same holds true of several flat-panel display technologies which can exceed the resolution of the CRT, as indicated in Table 1. While it is potentially true that resolution in excess of 1 arcminute may provide smoother graphics lines and avoid unwanted edge effects, the benefits of this approach need to be researched in terms of performance improvement rather than esthetic improvement.

TABLE 1. Comparison of Selected Parameters for Display Technologies

Technology	Resolution (Pixels/Inch)	Maximum Size (Inches)	Colors
CRT	360	30, diag.	Full range
Plasma Discharge	100	79, diag.	Orange
Electroluminescent	500	64 X 64	Red, Color in Develop.
Liquid Crystal	100	12 X 12	Any two
Electrochromic	unknown	unknown	Any two
Light Emitting Diode	320	10 X 10	Many Single

An added advantage occurs when pixel density becomes too great. Specifically, symbols or characters can be read most easily when the character block is on the order of 7 X 9 or 9 X 11 pixels. If the pixels become very small, then characters written in this fashion are too small for high legibility. The solution to this problem is to use two or more pixels for a single dot in the dot matrix character, thereby complicating both hardware and software design issues.

Uniformity

With new technologies and increased sizes come problems of nonuniformity of the image. In addition to the yield problems of uniform displays as the displays become larger, there is the general reluctance to replace a very expensive display if only a "few" pixels become nonusable. Thus, criteria for large area, small area, and individual pixel nonuniformities need to be established. To date, only a few research efforts have addressed this problem. For example, Abramson, Mason, and Snyder (1983) found that certain character fonts were more resistant to pixel loss than others, but that loss of only 2% of the pixels can have a significant effect on text reading time. It is important that acceptability and replacement criteria be defined experimentally for large area displays, particularly for information displays with nonredundant content.

Color

Although Christ (1975) amply demonstrated that color coding of displays could be beneficial, unimportant, or even detrimental, depending upon the application and task, most purchasers and system designers of today are adamant about the need for color displays, in spite of their increased cost, lowered reliability, and increased complexity.

Unfortunately, existing knowledge regarding the proper and improper use of color is disregarded. Recent evidence indicates that the "uniform color space" introduced in 1976 by the CIE is anything but perceptually uniform and does not predict well the effect of color differences upon visual task performance. In fact, a new color space based upon the CIE variables Y,u',v' has been shown to predict response speed much better than the previously used CIE spaces known as L*,u*,v* or L*,a*,b* (correlation of r = .93 vs. .83 and .81). The theoretical basis of this new metric has also been established, permitting designers or researchers to trade off color differences with luminance differences in predicting visual response speed to reading tasks (Lippert, 1984). Considerable additional research is warranted here, as the payoff in reducing display luminance in favor of suitable color combinations can be substantial, especially in high ambient illuminance environements.

In addition, the desire for high-resolution color is misguided because the visual system's spatial sensitivity to color is poorer than that to monochrome images (Figure 2). Thus, color display resolution need not approach that indicated above for monochrome displays.

Conclusions

Increasing size, resolution, and color content are strong trends in visual display technology. Yet the requirements for these "improvements" are doubtful in some applications and erroneous in others. Only through the development of a more quantitative and appropriate human performance data base will these technologies find their rightful place and not result in overkill or unneeded expense.

References

Abramson, S. R., Mason, L. H., and Snyder, H. L. The effects of display errors and font styles upon operator performance with a plasma panel. **Proceedings of the Human Factors Society**, 1983, 28-32.

Christ, R. E. Review and analysis of color coding research for visual displays. **Human Factors**, 1975, **17**, 542-570.

De Palma, J. J. and Lowry, E. M. Sine-wave response of the visual system. II. Sine-wave and square-wave contrast sensitivity. **Journal of the Optical Society of America**, 1962, **52**, 328-335.

Lippert, T. M. Color contrast effects for a simulated CRT head-up display. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, 1984.

Snyder, H. L. and Maddox, M. E. Information transfer from computer-generated, dot-matrix displays. Technical Report HFL-78-3, Virginia Polytechnic Institute and State University, 1978.

PANEL SESSION

A SYSTEMS APPROACH TO THE STUDY OF LEADERSHIP

Chair and Discussant: Edwin A. Fleishman, President
Advanced Research Resources Organization

Participants: William Haythorn (U.S. Army Research Institute)
Michael D. Mumford and Arthur L. Korotkin
(Advanced Research Resources Organization)

Background

Dr. Fleishman is internationally known for his pioneering research on leadership, development of taxonomies of human performance, and job analysis methodologies which link job task dimensions with ability, skill, and knowledge requirements. During this panel, Dr. Fleishman discusses the current research as it relates historically and developmentally to previous and other current approaches to the study of leadership.

Leadership Requirements on the Battlefield of the Future

William W. Haythorn

The Army, like most organizations, is faced with the need for gradual though continuous change as it adapts to changing environmental conditions. The leadership requirements in the Army of the 1990's may be different from those now existing. In particular, two predictable trends must be taken into account. First, the technology represented in the equipment used at all levels in the Army is rapidly becoming more complex, with substantial movement toward automation. There is a correlated increase in the volume and complexity of information that must be managed at each level. Second, the lethality of weapons is increasing rapidly. There is a correlated rapid increase in the total firepower controlled by small combat units.

A Theoretical Approach to Leadership

Michael D. Mumford
Kerry Yarkin-Levin
F. Mark Schemmer

The present paper reviews the literature on leadership, and notes the lack of a well founded theoretical conception of leadership which would provide a general, cross situational approach to leadership identification and development. Subsequently, it is suggested that a systems approach might be used to gain some understanding of leadership as it occurs in an organizational context. This leads to the hypothesis that formal leadership activities will always be focused on the attainment of certain goals specified in the leadership role, and thus will represent a specific form of problem solving activity. The literature supporting this hypothesis is reviewed. Moreover, the personal characteristics of the leader which would be likely to facilitate goal attainment and problem solving across situations are outlined, along with the impact of certain problem-specific skills. Finally, the implications of this approach for leadership identification and development will be discussed.

The movement toward higher technology applications may have impact on both the technical and conceptual skills required at various organizational levels. For example, the impact of automation may increase technical skill requirements for first level NCO supervisors without impacting on such requirements at considerably more senior levels. Conversely, the increasing diversity of resources which must be integrated into functional and combat effective units may well pose a requirement for significantly increased conceptual skills at more senior levels, particularly as these senior leaders are more and more likely to be "one step removed" from the actual mechanics of operation of a given system. Further, it might be postulated that interpersonal skill requirements will increase as senior leaders must increasingly often deal effectively with operators of technical systems which they themselves do not understand.

The increasing lethality of the conventional battlefield is rapidly producing a condition in which massed strength is an invitation to massive casualities. The "throw weight" of a division fielded by one of the major powers is now enormously greater than was the case in World War II. This leads to a visualization of the future battlefield as highly decentralized, or "distributed," in which smaller elements in greater number will be distributed in depth over a very large area, with semi-autonomous control over massive firepower. The requirements of command and control (leader functions) will be enormous.

Translating these requirements into a leadership research program will be discussed.

Leadership Job Dimensions and Competency Requirements for Commissioned and Non-Commissioned Officers

Arthur L. Korotkin
M. Reid Wallis
Ellen J. Eisner
Joanne Marshall-Mies

A current Army Research Institute-sponsored research project will be described which is directed at establishing a data base on which a systematic and progressive system of leadership development can be based. Such a system would utilize formal training, skills acquired through performance of assignments, and the sequence and pattern of assignments. Leadership is described in terms of a "systems" approach which attempts to integrate theoretical concepts and data from many sources into a coherent, pragmatic, and useful model of leadership.

The underlying conceptual structure of the model consists of two parallel hierarchical data sets, one job-based and the other person-based. On the one side, the elements of the job are described in varying levels of detail; and on the other side, the knowledges, skills, and abilities to perform the work are described. Both are tied to the overall or global goals of the Army and the more specific objectives which apply to the duty positions.

The methodologies for the development of Job Performance Dimensions (aggregates or clusters of tasks that reflect underlying similarities in work performed, and the identifivation of the knowledges, skills, and abilities required to perform the job) will be discussed. Data will be presented from the ratings of each JPD by non-commissioned (E-5 through E-8) and commissioned officers (01 through 06) with regard to time spent performing, criticality, amount of discretionary behavior, and time until results are seen. The results will be related to the projects theoretical foundation.

PANEL SESSION

PANEL SESSION: ANIMAL MODELS FOR THE ASSESSMENT OF HIGH RISK ENVIRONMENTS

SESSION CHAIR: Donald N. Farrer (USAF School of Aerospace Medicine)

PANELISTS: John H. McDonough (US Army Medical Research Institute of Chemical Defense)

Timothy F. Elsmore (Walter Reed Army Institute of Research)

Donald N. Farrer (USAF School of Aerospace Medicine)

John deLorge (Naval Aerospace Medical Research Laboratory)

Ben B. Morgan, Jr. (Old Dominion University)

PROCEEDINGS ENTRIES

"Animal models for the assessment of high risk environments" (John H. McDonough)

"Animal models of human performance: Structural and functional approaches to extrapolating from animal to man" (Timothy F. Elsmore)

"Animal models for the assessment of laser induced eye damage" (Donald N. Farrer)

"The assessment of nonionizing radiation hazards" (John deLorge)

"A review and evaluation of research concerning the performance effects of nuclear radiation" (Ben B. Morgan, Jr., *et al.*)

AD-P003 291

Animal Behavioral Testing in Medical Defense Against Chemical Agents

John H. McDonough, Jr., Ph.D., Major, MSC

U.S. Army Medical Research Institute of Chemical Defense
Aberdeen Proving Ground, MD 21010

Abstract

An overview is presented of the use of animal behavioral testing techniques with special application to the unique problems of medical chemical defense research. Representative animal behavioral tests to determine neurobehavioral effects of nerve agent exposure and performance decrements produced by antidote drugs are described.

There is an increasing potential threat for use of chemical warfare agents, in particular nerve agents, on the modern battlefield. The hazards of exposure to these agents and improved medical defense are currently the subject of intensive R&D efforts. Nerve agents are highly toxic organophosphorus compounds which exert their lethal effects by disruption of normal cholinergic neurotransmission in the central and peripheral nervous systems. Marked alterations in neurobehavioral function are characteristic of nonlethal exposure to these compounds and may persist for considerable periods of time following acute intoxication (Sidell, 1974). Many of the drugs used to treat poisoning by these agents are potent psychoactive compounds which have the capacity to alter sensory, motor and cognitive function. The development of improved drugs for the medical treatment of nerve agent casualties is highly reliant on the use of animal models to demonstrate both drug efficacy and safety. The prime determinant of drug efficacy is the ability to protect against the lethal effects of agent. The use of animal behavioral test models has extended this traditional toxicological approach to drug development. Behavioral testing models are used in two interrelated types of research studies: the assessment of the short- and long-term neurobehavioral consequences of exposure to nerve agent and the determination of side-effects of antidotes which degrade military performance. Testing methods *per se* for use in behavioral toxicology have recently been reviewed (Norton, 1982). Therefore, the rest of this paper gives examples of the application of several behavioral test models to each of these research problems.

Neurobehavioral Consequences of Nerve Agent Exposure

Sublethal exposure to nerve agents may produce long-lasting effects on brain function and behavior. Humans and animals exposed to the agent sarin demonstrated changes in electroencephalographic activity which persisted at least one year after exposure (Duffy, *et al.*, 1979). Exposure to the agent soman, under certain conditions, results in neuropathology (Petras, 1981; Lemercier, *et al.*, 1983) which is focused in limbic structures of the brain.

Casual observation of abnormal animal behavior after exposure was the impetus for the original Petras study. Recent work in our laboratory has been directed at determining the behavioral consequences of this nerve agent-induced brain damage.

Studies by physiological psychologists have shown that surgical destruction of limbic brain areas results in a constellation of behavioral deficits characterized by: changes in emotional behavior, memory function, and the ability to adapt responding to meet sudden changes in reinforcement contingencies. The choice of behavioral tests below was guided by this work. The subjects, rats, were first habituated to the laboratory, then randomly assigned to control (saline) or nerve agent (soman) exposure groups. After exposure all subjects were allowed to recover for between one to three weeks and then were subjected to behavioral testing. Emotional behavior was studied by rating responses to a series of normal handling procedures: resistance to capture, reaction to sensory strimuli, defecation, urination and vocalization. Memory function was tested using a one-trial passive avoidance procedure. In this test the animal received a single training trial where performance of a highly prepotent response was punished by shock. Memory for this experience was tested 24 hrs later and was considered intact if the animal did not respond. Acquisition of performance on a differential reinforcement of low rates (DRL) 20 sec operant schedule for food reward was used to test the ability to adapt performance to meet changes in reinforcement contingencies. After continuous reinforcement training the groups received 45 training sessions on the DRL schedule.

The behavior of nerve agent treated animals was markedly different from controls on all three tests of limbic system functions. Experimental subjects displayed an immediate and persistent increase in emotional reactivity. They were hyperreactive to normal handling and violently resisted being held. Increased emotionality scores were positively correlated to the intensity of the acute symptoms of poisoning. On the test of memory function 14 of 29 exposed animals failed to display normal behavior even during the training trial, so memory function could not be assessed in these subjects. Of the remaining agent-exposed subjects (N=15) 33% showed no evidence of remembering the training experience. All normal control subjects (N=36) showed perfect memory for the aversive experience. On the DRL 20 sec operant test, normal rats displayed a 140% improvement in the number of rewards earned over the 45 sessions of training while nerve agent exposed subjects improved only 50%. In addition, agent-exposed animals had not developed the response distribution pattern typically engendered by this schedule even after 45 sessions while controls had shown clear development of this pattern by session 15. The results of these tests show that a single exposure to a high dose of nerve agent can produce long-lasting behavioral abnormalities in test subjects. The behavioral deficits are indicative of damage to limbic system structures. Further testing is needed to define other possible deficits, and perhaps more importantly what capabilities are retained in exposed subjects. Other work is currently in progress to test whether higher species are susceptible to this effect of agent, the mechanisms by which the damage is produced, and how it can be attenuated with antidote drugs.

Behavioral Side-Effects of Antidote Drugs

Many drugs have the capacity to produce undesirable effects which are incidental to their major therapeutic indication, such as the drowsiness produced by antihistamines. The side-effects of drugs can have significant adverse impact on the performance of everyday tasks. Many drugs developed as antidotes for nerve agent poisoning are potent psychoactive compounds and an accurate knowledge of the types of behavioral functions that are degraded and the duration of these effects must be known. Laboratory studies, especially with nonhuman primates, can model various functional aspects of behavior which contribute to successful performance of many tasks. Results of such studies provide a method to generalize and predict the relative effect of a drug over a broad range of conditions, and in the case of a novel drug, guide the selection of behavioral tests to be performed in human clinical trials. Perhaps the best example of this use of animal behavioral models are the series of tests used to study the effects of anticholinergic antidote drugs. Although there was considerable literature on the effects of anticholinergics as a class of drugs in both humans and animals, prior to 1979, there was sparse information as to the behavioral effects of anticholinergics in nonhuman primates. Especially lacking was any data on dose-effect relationships or time-course of action. Knowledge of this information was needed in order to relate behavioral performance decrements to the drug doses demonstrated to be antidotally effective against nerve agent poisoning in toxicological studies (McDonough & Penetar, 1982).

Three behavioral tests were used to determine the effects of two anticholinergic drugs, atropine and benactyzine, on nonhuman primate performance. Details of the experimental tasks and procedures have been published (Bennett, et al., 1981; Farrer, et al., 1982; McDonough, 1982; Penetar & McDonough, 1983). The general procedure in all tests was to first train the subjects to stable levels of performance. Drugs were then given in standardized doses, by the same route, using counterbalancing and Latin-Square design procedures to control for order effects. The tasks choosen were: continuous equilibrium tracking with concurrent multiple response, DRL 28 sec operant schedule, and delayed match-to-sample (DMS). Performance in the DRL and DMS tasks were food motivated, while that in the equilibrium-multiple response task was shock motivated. The equilibrium-multiple response task provided the primary data on dose-response and time-course effects of the drugs while the DRL and DMS tasks assessed vigilance performance and short-term memory function, respectively.

The results of the equilibrium-multiple response task showed these two anticholinergic drugs to have distinctly different potencies and time-courses of action. Atropine produced effects at doses $\geq$ 0.10 mg/kg and performance disruption was near maximal at doses > 0.40 mg/kg. In addition there was a delay in onset of effect > 30 min and peak effects lasted almost 6 hr. In contrast, benactyzine began to produce performance decrements in doses $\geq$ 0.50 mg/kg and much larger increases in dose were needed to further degrade performance. This drug also had a very rapid onset of effect ($\leq$ 15 min) but lasted only a short period of time (2-3 hr).

Both atropine and benactyzine degraded performance on the DRL 28 sec task at doses that did not produce reliable effects on response rate. These

effects occurred in the effective dose range observed in the previous study. Regression analysis of the performance decrement dose-effect curves indicated that benactyzine may act by a different mechanism than atropine since it was both less potent and had a significantly shallower slope. Drugs acting by similar mechanisms but of differing potencies should have offset yet parallel slopes.

The results of the DMS study demonstrated that only atropine affected short-term memory function. Benactyzine had only a general disruptive effect on performance. The fact that short-term memory was specifically affected by atropine was demonstrated by a significant dose by delay interaction in error rates. At the higher doses of drug, matching performance became progressively worse the longer an animal had to remember the sample stimulus. At the highest drug dose matching performance was at chance levels when the animals had to remember the sample stimulus for greater than 8 sec.

From these results good predictions could be made as to the effects of these two drugs on human performance. Both drugs would reduce vigilance and the ability to concentrate. Time to complete tasks would be longer due to increased errors. Tasks requiring a recent memory component would be seriously degraded by atropine. At equivalent doses performance would be at greater risk and for longer periods from atropine, while benactyzine would degrade performance more rapidly. Information such as this would be especially valuable to psychologists planning to study the effects of these drugs on cognitive functions during clinical trials.

The experiments reported here were conducted according to the Guide for Care and Use of Laboratory Animals (1978) as prepared by the Committee on Care and Use of Laboratory Animals, National Research Council, DHEW PUB. NO. (NIH) 80-23. The opinions and assertations contained herein are the private views of the author and are not to be construed as reflecting the views of the Department of the Army or Department of Defense.The author gratefully acknowledges CPT M. Mays, Ph.D., CPT H. Modrow, Ph.D., and R. Smith, Ph.D., for allowing use of their data and their support in this manuscript, and to the fine technicians who assisted these studies.

References

Bennett, C.T., Lof, N.E., Farrer, D.N., & Mattsson, J.L. Equilibrium performance changes produced by atropine in M. mulatta and M. fascicularis. USAFSAM Tech. Rep.-81-29 (Nov 1981), Brooks AFB, TX.

Duffy, F., Burchfiel, J., Bartels, P., Gaon, M., & Sim, V. Long-term effects of an organophosphate upon the human electroencephalogram. Toxicology & Applied Pharmacology, 1979, 47, 161-176.

Farrer, D.N., Yochmowitz, M., Mattsson, J., Lof, N., & Bennett, C. Effects of benactyzine on an equilibrium and multiple response task in rhesus monkeys. Pharmacology Biochemistry & Behavior, 1982, 16, 605-609.

Lemercier, G., Carpentier, P., Sentenac-Roumanou, & Morelis, P. Histological and histochemical changes in the central nervous system of rat poisoned by an irreversible anticholinesterase organophosphorus compound. Acta Neuropathologica, 1983, 61, 123-129.

McDonough, J.H. Effects of anticholinergic drugs on DRL performance of rhesus monkeys. Pharmacology Biochemistry & Behavior, 1982, 17, 85-90.

McDonough, J.H., & Penetar, D.M. The effects of cholinergic blocking agents and anticholinesterase compounds on memory, learning and performance. In M. Spiegelstein & A. Levy (Eds.), Behavioral models and the analysis of drug action. Amsterdam: Elsevier, 1982.

Norton, S. Methods in behavioral toxicology. In A.W. Hayes (Ed.), Principles and methods of toxicology. New York: Raven Press, 1982.

Penetar, D.M., & McDonough, J.H. Effects of cholinergic drugs on delayed match-to-sample performance of rhesus monkeys. Pharmacology Biochemistry & Behavior, 1983, 19, 963-967.

Petras, J.M. Soman neurotoxicity. Fundamental & Applied Toxicology, 1981, 1, 242.

Sidell, F.R. Soman and sarin: clinical manifestations and treatment of accidental poisoning by organophosphates. Clinical Toxicology, 1974, 7, 1-17.

AD-P003 292

Animal models of human performance: Structural and functional approaches to extrapolating from animal to man[1,2]

Timothy F. Elsmore
Walter Reed Army Institute of Research
Washington, DC 20307

Abstract

Research with non-human animals is performed to solve human problems. Two different means of extrapolating from behavioral experiments with animals to human behavior are described. The structural approach emphasizes topographical similarity between the animal test situation and the human situation of interest. The functional approach emphasizes behavioral functions and controlling variables, and minimizes the importance of topographical similarity. Two experiments are presented to illustrate the functional approach, one dealing with the effects of chronic marihuana use upon performance, and another dealing with the issue of memory deficits produced by chemical warfare antidotes.

Introduction

The primary reason we do research with animals is to solve human problems. **Animal models** of human performance are those behavioral experiments we perform with animals in an attempt to gather information that is relevant to significant human situations. The question is, how valid are these models? In other words, what conclusions can we draw from our animal models that permit us to more effectively deal with the human situation? When we are addressing problems of considerable practical importance such as those related to performance decrements produced by compounds used for defense against chemical weapons, it is critical to know if the performances we examine in our animals have anything to do with our target organism, the soldier.

Animal models of human performance may be conceptualized as lying along a continuum. One end of this continuum may be characterized as **structural**, and the other end as **functional**. I will now briefly define these end points of the animal model continuum, and then give some examples of experiments illustrating the functional approach. My objective is to illustrate how functional models may provide us with information that is more broadly applicable than that provided by structural models.

Structural models. Structural models are attempts to simulate human tasks with animals. These models typically involve complex stimulus displays and require the animal to perform complex response topographies, and focus on **what** the animal is doing. Structural models are designed to have great face validity, such that we can easily identify with what the animal appears to be doing. Unfortunately, most human situations are exceedingly complex, and our attempts to simulate these situations with non-human animals tend to create situations in which some variables may not be under

---

1. The views of the author do not purport to reflect the position of the Department of the Army or the Department of Defense, (para 4-3, AR 360-5).

2. In conducting the research described in this report, the investigators adhered to the 'Guide for the Care and Use of Laboratory Animals', as promulgated by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council.

good experimental control, and to make interpretation of results difficult. Structural models tempt us to anthropomorphize; that is, to interpret the behavior of our animals as if they were doing human-like things in the same way and for the same reasons humans do them, when that may not at all be the case. Attempts to reduce this complexity necessarily also reduce the face validity of the model. A practical problem with the structural approach is that animals, even clever monkeys, can only be trained to perform a very limited subset of the tasks required of soldiers in the field, thus the conclusions that can be drawn from these experiments are accordingly limited.

Examples of structural models are provided by the many recent experiments that have been performed which attempted to teach chimpanzees verbal behavior of various sorts. Initially, these studies appeared to be successful in generating human-like verbal behavior in chimps, but more careful analysis has showed significant differences between chimpanzee 'verbal behavior' and human verbal behavior (e.g. Savage-Rumbaugh, Rumbaugh, Smith, and Lawson, 1980). While it would be highly desirable to have an animal model of human verbal behavior, it seems unlikely that this will come about in the near future. In any case, the investment of resources required to train such skills in non-human animals would make the use of such a model impractical for any realistic applied testing program.

Functional models. An alternative way to proceed is to conceptualize performance in functional terms. This involves attempting to identify the variables that are important in producing critical aspects of military performance and the behavioral functions that are components of these performances, and investigating these, rather than concentrating on the structural or topographical aspects of particular military tasks. Functional models focus on **why** the animal does what it does. Examples of this approach would be studies on memory, learning, and performance under different schedules of reinforcement. Obviously, there are hazards in this approach as well. Our interpretation of what is involved in producing normal human performance in given situations may be in error. Our implementation of animal experiments designed to investigate the 'same' variables may also be in error. There will undoubtedly be parametric problems with motivation and stimulus intensity, as well as other variables. Despite these problems, the benefits of a functional approach may be great. Experimental results should be more general, since a variety of different human tasks presumably involve the same functions. A greater degree of experimental control may be possible than in studies designed to simulate 'human' situations, and from a practical point of view, it usually will be possible to use less elaborate equipment, training procedures, and less expensive animals. I would now like to illustrate the use of a functional approach in two different experiments, which were done for quite different reasons.

Marihuana and timing behavior in monkeys

In the early seventies, it became apparent that the U.S. military had a significant problem with drug abuse, particularly with soldiers serving in Southeast Asia, and our group at Walter Reed was tasked with doing some research in this area. One question that was asked was, 'Can soldiers who chronically smoke marihuana or hashish continue to perform their jobs?' At the time, there was a great deal of folklore surrounding marihuana, and even a few controlled laboratory studies with both human and non-human animals, but no reliable information on the effects of chronic use of this substance. One approach to answering this question would be to attempt to get permission to chronically administer marihuana or its active ingredient W-9 tetrahydrocannabinol (W-9 THC) to human volunteers, and evaluate the effects of such treatment on a variety of behavioral functions. You are probably all aware of the difficulties of such a project. Therefore we

elected to carry out this study with monkeys (Elsmore, 1976), and to focus on timing behavior since the available literature suggested that marihuana produces large changes in the ability of animal and human subjects to judge the passage of time (Conrad, Elsmore, and Sodetz, 1972; Hollister, Richards, and Gillespie, 1968).

Two rhesus monkeys were the subjects in this experiment. They each lived for the duration of the experiment in a cage with a lever protruding from one wall. During experimental sessions a small pilot light over the lever was illuminated and the animals could earn food pellets by pressing on the lever. The way in which the pellets could be earned, or schedule of reinforcement, depended on an auditory stimulus which was also present. When the stimulus was white noise, a fixed interval 2 minute (FI 2 min) schedule was in effect. Under this condition, the first press following 2 min from the onset of the white noise produced a food pellet and changed the auditory stimulus to a 1500 Hz tone. When the tone was on, the animal was required to **withhold** responding for 2 min before a lever press produced a pellet. Any press prior to 2 minutes since the start of the tone or since the last response would reset the 2 min timer (DRL 2 min). These two conditions alternated throughout the experimental session. The important thing to note is that both the FI and DRL schedules of reinforcement required the animal to time a 2 min interval for efficient performance. The animals were trained for about three months under this procedure until their performance was quite efficient, and appropriate to the particular schedule in effect. They were then treated with a 7 mg/kg oral dose of Δ-9 THC for 40 consecutive days, two hours prior to the experimental session.

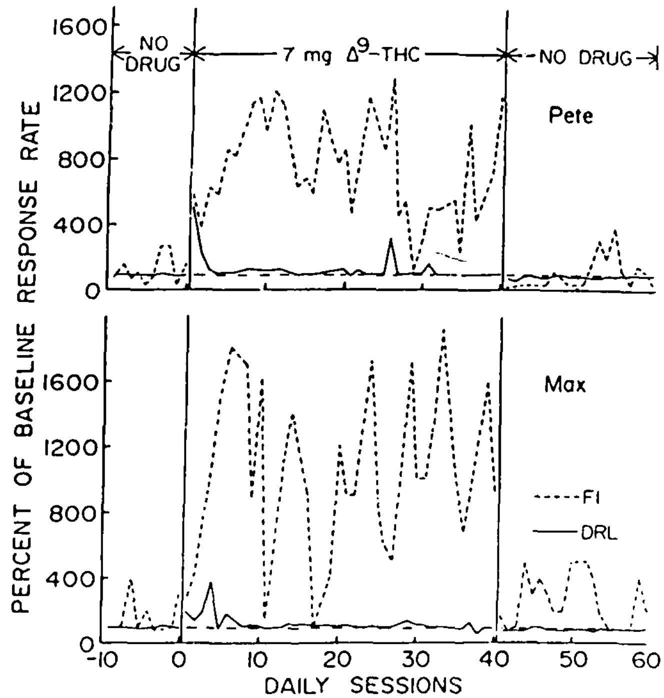


Figure 1. Effects of daily administration of Δ-9 THC (marihuana) upon food-reinforced responding under two different schedules of reinforcement.

Figure 1 shows the major results of this experiment. Each frame is for a single animal. When Δ-9 THC was administered, shown by the first vertical line, both animals showed an immediate increase in responding relative to baseline levels. With chronic administration, however, responding rapidly returned to normal levels for the DRL

schedule, but not for the FI schedule. Why did the animals' performance recover only in the DRL component? The reason is that increased responding under the DRL schedule cost the animals pellets, whereas increased responding under the FI schedule only decreased the efficiency of their performance and did not affect the delivery of food pellets. When the drug had a deleterious effect on food intake, the animals adapted their performance to accommodate the drug effect. When it did not make much difference, no adaptation occurred. These results suggest that we must focus on the consequences of a drug-induced performance decrement to answer the question, 'Can soldiers who chronically smoke marijuana or hashish continue to perform their jobs.'

Effects of atropine on memory in the mouse

My second example involves a very different animal, the mouse, and a very different problem. The chemicals used in nerve agents, and the antidotes to these agents act strongly on the neurotransmitter acetylcholine, which has also been implicated in the neurochemistry of memory (Drachman, 1977). It is natural, therefore, to look for an animal model of memory which may be used to investigate the effects of CW antidotes upon memory, since it is important to know what effects these agents might have upon military tasks involving memory. Serious effects must necessarily be taken into account in the creation of doctrine regarding the use of antidotes.

One relatively cost-effective model of memory in animals is the radial maze (Olton and Samuelson, 1976). The maze we used (Levy, Kluge, and Elsmore, 1983) had a hexagonal central platform with six radial arms. The central platform was 15 cm in diameter, and each arm was 30 cm long with a small plastic food cup at the end. Each animal was given one trial each day in which three of the arms (always the same three for each animal) contained food, and three did not. Within 20 such trials, the animals had learned to enter only those arms containing food, and to avoid those which never contained food. Although the performance of all mice was not perfect, it was significantly better than chance. The animals were then treated with atropine sulfate in doses ranging from 1 to 6 mg/kg, one hour prior to their daily sessions. Atropine was given only once per week.

Table 1

Atropine Effects 1-hr Postinjection[a]

Drug/ dose (mg/kg)	Time per choice (seconds) (TPC)	Reference memory errors (RM)	Working memory errors (WM)	Working memory incorrect (WMI)	Repeated errors (RE)
AT/1[b]	8.33	1.10	0.33	0.05	0.38
AT/2	14.69	1.50	0.95	0.23	1.18
AT/4	17.15	1.91	1.00	0.64	1.64
AT/6	19.87	1.18	1.09	0.41	1.50
Saline	7.07	1.71	0.18	0.05	0.23

[a] n = 22, mean values are given.
[b] AT = Atropine sulfate.

Results are shown in Table 1. The first four rows of this table show the average effects of different atropine doses upon different measures of performance, and the last row shows the control values for the same measures. The dose-related effect of atropine

is shown by the effects on time per choice (TPC) in the first column, which increased by a factor of 2.8 at the highest drug dose. The most interesting effect, however, is seen in a comparison of reference memory and working memory errors in the next two columns. Reference memory errors, which are entrances to arms of the maze which were never baited with food, showed no significant increase with drug dose. Working memory errors, which were operationally defined as reentering arms from which the food had already been taken, increased dramatically to more than 6 times control values at the highest dose. Thus, atropine appears in this situation to affect working memory or short-term memory, but not to affect reference or long-term memory. Thus, we might assume that under the influence of atropine, soldiers might remember what they are supposed to do in particular situations, but might have difficulty remembering whether they have completed a particular task. This is obviously a rather large leap from mouse to man, but this study provides some suggestions as to what types of deficits might be profitably investigated in more expensive experiments with non-human and human primates.

Conclusions

I have tried to demonstrate, by the use of two very different experiments, a functional approach to extrapolating from behavioral data obtained with non-human animals to human situations of practical importance. Many assumptions are required to make this leap, but we must begin somewhere. The simplicity and experimental control inherent in the functional approach allows us to carefully study particular variables in relative isolation, and to make conclusions that are valid within their own experimental contexts. The structural approach, by virtue of its complexity, makes it more difficult to arrive at conclusions that are valid even within the limited confines of the model, and easier to make unwarranted inferences regarding the implications of experimental results for human behavior.

Animal research is necessary to guide research with humans, and to investigate variables that cannot be investigated with human subjects. Thoughtful judgements about our choices of animal models must be made to efficiently and economically use our scarce research resources. The distinction between structural and functional models can be a help in this regard. Although as indicated earlier we must think in terms of a structural -- functional continuum, I believe it is a useful distinction which can serve to organize our thinking in the planning and execution of research programs with non-human species. While arguments can be made for both structural and functional approaches to animal research, we must be aware of the pitfalls of whatever approach we choose to avoid making conclusions that are not justified by our data.

References

Conrad, D.C., Elsmore, T.F. and Sodetz, F.J. Δ-9 Tetrahydrocannabinol: Dose-related effects on timing behavior in chimpanzee. Science, 1972, 175, 547-550.

Drachman, D.A. Memory and cognitive function in man: Does the cholinergic system have a specific role? Neurology, 1977, 27, 783-790.

Elsmore, T.F. The role of reinforcement loss in tolerance to chronic Δ-9 tetrahydrocannabinol effects on operant behavior of rhesus monkeys. Pharmacology, Biochemistry, and Behavior, 1976, 5, 123-128.

Hollister, L.E., Richards, R.K. and Gillespie, H.K. Comparison of tetrahydrocannabinol and synhexyl in man. Clinical Pharmacology and Therapeutics, 1968, 9, 783-791.

Levy, A., Kluge, P.B. and Elsmore, T.F. Radial arm maze performance of mice: Acquisition and atropine effects. Behavioral and Neural Biology, 1983, in press.

Olton, D.A. and Samuelson, R.J. Remembrance of places passed: Spatial memory in rats. Journal of Experimental Psychology: Animal Behavior Processes, 1976, 2, 97-116.

Savage-Rumbaugh, E.S., Rumbaugh, D.M., Smith, S.T. and Lawson, L. Reference: The essential linguist. Science, 1980, 210, 922-925.

AD-P003 293

Animal Models for the Assessment of Laser Induced Eye Damage

Donald N. Farrer
United States Air Force School of Aerospace Medicine

Abstract

This review paper summarizes the experimental findings of laser induced eye damage experiments. This report describes experiments in which dependent variables range from histology and minimum ophthalmoscopically visible damage to the assessment of functional vision. Laser induced retinal lesions which are very small do not measurably alter functional vision although the fovea is permanently damaged. The assessment of functional vision and quantifiable losses depend upon operant behavioral methods to expose subjective visual acuity to examination in the laboratory animal. Quantifiable electrophysiological techniques offer promise for objective functional vision assessment alternatives.

The first reports of eye damage resulting from an intense light source were recorded by Galen and described as an "eclipse blindness." Later, Galileo's eye damage was attributed to his observations of the sun, but only in modern times, the phenomenon has been described as thermal injury to the pigment epithelium and choroid resulting from the image of the sun focused on the retina.

Technological advances have resulted in several intense optical devices such as nuclear fireballs and a family of laser beams which pose serious threats to functional vision required for mission accomplishment by military personnel. Also, space exploration, civilian supersonic transportation, industrial laser applications, as well as other intense optical sources, have been developed which introduce the possibility of exposing appreciable elements of the civilian population to eye damage in the form of retinal burns.

In an effort to provide safety criteria for people who must work in these hazardous environments, several research workers have attempted to define a retinal burn threshold in terms of the power density and exposure time which produces permanent changes in the retina. Many early experiments of this type were conducted with animals in experimental designs which used nuclear devices detonated in the atmosphere, and blindness was inferred from the magnitude of the retinal lesion. Additional data were obtained with experiments on animals exposed to white light sources such as the xenon lamp and various laser devices to simulate operational hazards to vision. Several investigators felt that the anatomical differences between the eyes of non-primates and men were significant, and the rhesus monkey was selected for the further detailed study of retinal burns.

The macaques have a visual system which is very similar to that of man. The macaque retina contains both rods and cones, a distinct macula, and an all cone fovea. This anatomical similarity is supported by the additional fact that the average subjective visual acuity of the young adult rhesus monkey is identical to the average of young men (Weinstein and Grether, 1940). Within a few years, the rhesus monkey became a very popular research subject for studies of eye hazards, in the form of retinal burns, using several intense optical sources. The dependent variable in the majority of these experiments concerned the determination of an opthalmoscopically observable lesion in the fundus, and the majority of investigators concluded their research with histological studies using light/or electron microscopy which attempted to discern damage to the retina at energy density levels below those required to produce a lesion which is visible with opthalmoscopy. Using histological techniques, damage to the retina was demonstrated at energy densities too low to produce visible retinal burns. Thus, a substantial amount of data have been obtained to indicate the potential danger of permanent eye damage caused by exposure to intense optical sources.

The generally accepted safe exposure limit for human experimentation with intense optical systems is less than 10% of the energy which is required to produce a minimum opthalmoscopically observable lesion in an experimental animal such as a rabbit or a monkey. These data are obtained by measuring the total energy entering the eye, calculating the image diameter on the retina, and the exposure time necessary to produce a barely opthalmoscopically visible lesion on the retina. Usually, such a lesion is minimal and a lapse of time is necessary after exposure before it becomes observable. Many investigators look for the presence of a lesion five minutes after exposure as a practical time limit, while others prefer to establish thresholds on the basis of a 24 hour postexposure determination.

Jones and McCartney (1966), published photographs showing some photoreceptor degeneration following laser exposure at energy densities below the levels required to produce an opthalmoscopically observable lesion. These findings established the need for experiments of the assessment of functional vision.

Primates were again selected as the subject most suitable for functional vision assessment. Yarckzower and Wolbarsht (1966), trained a stumptailed macaque to accurately discriminate between two visual targets which subtended 1.4 minutes of arc (20/28 Snellen), then completely destroyed the macula of the anesthetized subject and two days later retested the ability of the subject to discriminate between visual targets subtending small visual angles. Following foveal destruction, the smallest target which the subject could discriminate subtended 10.0 minutes of arc (20/200 Snellen). This experiment was the first to demonstrate the discrimination technique with nonhuman primates for the assessment of changes in functional vision following thermal injury to the retina.

A comprehensive research program was initiated to study the functional vision changes to be expected from a variety of intense optical sources, including xenon lamp flashes, ruby and neodymium laser pulses, and at energy densities above and below opthalmoscopically observable damage, and in time courses ranging from immediate assessment to one year postexposure studies in our laboratory. Functional vision changes were studied following brain x-irradiation (Graham, Farrer and Carsten, 1969), and acceleration stress (Leverette and Farrer, 1970). The first research goal of this program concerned the determination of subjective visual acuity in rhesus monkeys. The variables to be considered were: lesion size, duration of visual impairment and viewing distance (3 feet vs. 20 feet), as well as discrimination stimuli and tasks.

The methods employed for the determination of subjective visual acuity are described by Farrer and Graham (1967), and the details of the apparatus are described by Graham, Farrer, Crook, and Garcia (1970).

The lesion size studies produced remarkable findings given the earlier Yzarczower and Wolbarsht data point. We found that acuity did not change from baseline levels with small (e.g., 50 microns to 100 microns in diameter) lesions placed directly in the center of the fovea. Indeed, much larger lesions (e.g., 300 microns in diameter) in the fovea which degraded vision to approximately 20/100 Snellen was observed to recover to better than 20/20 Snellen within 90 to 120 days postexposure. This recovery of functional vision was not accompanied by total retinal repair. The foveal lesion was still opthalmoscopically observable, but the functional vision had improved. Thus, it was evident that visual capability could not be accurately predicted from ophthalmic observations alone.

The difficult question of immediate versus delayed effects remained. Using anesthetized subjects allowed very precisely located lesion placement and lesion size studies, but such preparations did not answer the questions about functional vision at the moment of exposure. These studies required more restraint (fixed eye position) monocular viewing tubes and Purkinji image reflections to accurately stimulate the fovea with small laser pulses while the subject performed the operant discrimination tasks (Farrer, 1973), and the results reflected surprisingly good vision remained after high intensity monochromatic flashes. The first exposure caused a startle reaction which impaired performance, and additional training was required to adapt the subjects to continue to perform the visual acuity task following laser exposures at intensities above the OSHA and AFOSH safety standards, but below lesion intensities. No permanent visual impairment (visual acuity) has been observed from any one of the subjects unless the foveal lesion was ophthalmoscopically observable lesion was larger than 300 microns. This finding is not to imply that laser radiation is safe at energy densities above the American National Standards Institute safety threshold standards. Indeed, Zwick (1978), reported color vision impairment following very low light levels from a continuous wave Argon laser with two subjects. Similar findings were first reported by Harwerth and Sperling (1971). These minimal damage levels clearly support the safety standards as currently enforced, but exposures which exceed these sure-safe levels should not be considered as causing immediate and permanent blindness.

A series of studies was conducted by Callin and his coworkers to examine visual compensatory tracking performance as a function of intense light sources. In their first study, Callin, Devine and Garcia, 1981a, they demonstrated that the duration of flashblindness (e.g., recovery of useful vision following an induced after image) was not significantly different in human and rhesus monkeys. Given these white light data obtained at low and safe levels of illumination, the next study (Callin, Devine and Garcia, 1981b) involved exposing the animal subjects to higher energy density laser sources. Flashblindness was not produced without producing lesions. This finding was explained in the third study (Callin, Devine, and Garcia, 1981c) as a function of illumination spot size. The collimated laser source was focused on the retina within the accuracy as permitted by the optics of the subjects eye (refractive error), and as in the retinal burn data reported earlier, the redundancy of the fundus is sufficient to resolve the relatively large targets (greater than one minute of arc). Thus, these studies yielded data which indicate that damage to the fundus occurs at energy densities lower than the laser energy required to impair useful and normal vision.

The new frontiers of visual science such as contrast sensitivity and photo evoked potentials offer laboratory alternatives to operant methods for the assessment of visual decrements. Previc (1983), has shown electrophysiological changes as a function of laser flashes which correlate well with after images. These methods used together can enhance our ability to assess functional vision capability in hazardous laser environments.

<u>References</u>

Callin, G. D., J. V. Devine, and P. Garcia. Visual compensatory tracking performance after exposure to flashblinding pulses: I. Comparison of human and rhesus monkey subjects. USAFSAM TR-81-3 (1981a).

Callin, G. D., J. V. Devine, and P. Garcia. Visual compensatory tracking performance after exposure to flashblinding pulses: II. Subdamage threshold laser irradiation of rhesus monkey subjects. USAFSAM TR-81-7 (1981b).

Callin, G. D., J. V. Devine, and P. Garcia. Visual compensatory tracking performance after exposure to flashblinding pulses: III. The importance of retinal image diameter. USAFSAM TR-81-8 (1981c).

Farrer, D. N. and E. S. Graham. Visual acuity in monkeys a monocular and binoocular subjective technique. Vision Res, Vol 7, pp. 743-747 (1967).

Farrer, D. N. Studies in biomedical research. In NonHuman Primates and Biomedical Research. G. Bourne (Ed.) Academic Press, NY (1973).

Graham, E. S., D. N. Farrer, and A. L. Carsten. Blindness in the rhesus monkey as a delayed effect of radiation exposure of the occipital lobes. Rad Res,Vol 39 (1969).

Graham, E. S., D. N.Farrer, G.H. Crook, P. V. Garcia. A self-adjustment procedure for measuring the visual acuity of rhesus monkeys. Behav Res Meth and Instru, Vol 2(6), pp. 301-305 (1970).

Leverette, S., and D. N. Farrer. Personal communication of unpublished data (1970).

Harwerth, R. S. and H. G. Sperling. Prolonged color blindness induced by intense spectral lights in rhesus monkeys. Science, Vol 174, pp. 520-523 (1971).

Jones, A. E. and A. J. McCartney. Ruby laser effects on the monkey eye. Invest Ophthal, Vol 5, pp. 474-483 (1966).

Previc, F. Personal communication of unpublished data (1983).

Weinstein, B. and W. F. Grether. A comparison of visual acuity in the rhesus monkey and man. J Comp Psychol, Vol 30, pp. 187-195 (1940).

Yarckzower, M., and M. L. Wolbarsht. Visual acuity in the stumptailed macaque. Science, Vol 152, pp. 1392-1393 (1966).

Zwick, H. Low-level laser light effects. Spie, Vol 162, pp. 112-118 (1978).

THE ASSESSMENT OF NONIONIZING RADIATION HAZARDS

John de Lorge
Naval Aerospace Medical Research Laboratory

Abstract

Naval personnel frequently occupy environments susceptible to microwave energy from weapons systems and communication devices. Several scientific reports have indicated that behavior can be modified by relatively low levels of microwave energy, but the detrimental or beneficial nature of these behavioral effects are not immediately apparent. A series of experiments employing a behavior whose affective nature could be assessed and which was analogous to human behavior was conducted with several species of animals. Rats, squirrel monkeys, and rhesus monkeys trained on operant tasks were exposed to microwave radiation and produced data suggestive of a possible extrapolation to man in similar situations. The results show that behavioral changes are related to increases in colonic temperature. In monkeys the average increase in colonic temperature associated with changes in response rate was 1 °C. Response rate did not change in the absence of concomitant temperature increases.

Introduction

Background

A major problem in evaluating the effects of any physical agent that may be potentially hazardous to man is that man himself is not likely to be the experimental subject. Selecting the most appropriate organism for testing that agent and applying the results to man then becomes an often controversial task. The traditional subject for testing the effects of nonionizing radiation, particularly radio frequency (RF) radiation, is the rat. However, generalizations drawn from rats and applied to man are especially prone to error when dealing with RF radiation. The effects of RF radiation are so dependent on the geometry and body mass of an animal that it may be impossible to position man in the same orientation to radiating sources as one does a rat. Hence, animals more similar to man are needed. One such animal is the squirrel monkey, Saimiri sciureus; another is the rhesus monkey, Macaca mulatta. These animals along with the rat provide animals of three distinctly different body masses that relate to one another on an almost logarithmic scale thereby providing one dimension for extrapolation to man.

Another problem in evaluating biological effects is choice of the dependent variable. Psychologists are continuously tasked with generalizing from behavior in animals to behavior in man. One solution is to train animals to perform a task in the same manner as required of humans. A general class of behaviors that encompass such performance is designated as operant. The present series of experiments (de Lorge, 1976; 1979; in press; de Lorge and Ezell, 1980) utilized operant behavioral changes as the dependent variable against which the effects of RF radiation were measured. In most

cases an observing response task was chosen because of its similarity to vigilance behavior in man (Holland, 1957). In other instances the results from animals performing different operant tasks (Sanza and de Lorge, 1977; Nelson, 1978, Knepton and de Lorge, 1983) will be reported.

So that the interaction of body mass and frequency of RF radiation could be examined, uniquely different frequencies were utilized in these experiments. The frequencies of 225 MHz, 1.3 GHz, 2.45 GHz, and 5.7 GHz were also chosen because of availability and because they represented frequencies in use by the Navy. The rhesus monkey of the size used in these studies has a resonant RF frequency of about 225 MHz; the other three frequencies were not only above the resonant frequencies of the rhesus but also of the rat and squirrel monkey.

The general approach of these studies was to train animals on operant tasks until stable response rates were obtained, and then expose the animals to RF radiation during their work sessions. This approach demonstrated that in all of the animals an operant response was well maintained during hour-long sessions until an animal's colonic temperature reached or surpassed a 1 °C increase above its baseline.

Method

Subjects

Rats obtained from the Charles River colonies, squirrel monkeys, *Saimiri sciureus*, obtained from Columbia, South America, and rhesus monkeys, *Macaca mulatta*, bred in our own colony were subjects. Body masses of the rats were 300-400 g, the squirrel monkeys averaged 700 g, and the rhesus monkeys ranged between 4.3 and 5.7 kg. The animals were food deprived and trained at approximately 85% of their free-feeding body mass. Body mass during sham and exposure sessions was typically somewhat higher; for example, the rhesus monkeys were maintained at 92% of their free-feeding body mass during their experimental sessions. The animals normally obtained their daily food during their sessions and were supplemented in their home cage. Water was continuously available in the home cages.

Apparatus

Four anechoic chambers differing only in basic dimensions were used for radiation exposure. The chambers were lined with pyramidal absorber, cooled with air conditioners and ventilation fans, and each was equipped with a television camera. Devices for presenting auditory and visual stimuli, food reinforcement, chamber illumination, and monitoring temperature and humidity were also located in the chambers.

RF radiation was generated by various military and commercial radar sets at .225, 1.28, 2.45 and 5.62 GHz. The .225 and 2.45 GHz sources were continous wave whereas the 1.28 and 5.62 GHz sources were pulsed at different repetition rates. Both custom-made and standard gain horns were used. In the monkey experiments the front surface of the upright seated animal was usually irradiated, but dorsal irradiation was used with one squirrel monkey

study and a rat experiment. The rats were irradiated on their right side in other studies. Colonic temperatures were obtained from the monkeys during their experimental sessions.

Procedures

Animals were initially trained in Plexiglas devices and then transferred to Styrofoam boxes in the case of the rats (Sanza and de Lorge, 1977) and chairs in the cases of the monkeys (Reno and de Lorge, 1977).

The typical operant task required an animal to respond on one of two different levers to produce stimuli or food pellets. Stimuli were randomly presented depending upon lever response rate and when the stimulus corresponding to availability of food appeared a response on the second lever produced food. The schedule of reinforcement, number of manipulanda, stimulus modality and quantity varied in different experiments. Food was the only reinforcer used and water was unavailable during the experimental sessions which lasted from 40 minutes to 2 hours. The rate of responding on the lever producing the food signal was the prime measure of performance.

In all of these experiments session-to-session response rate stability was established prior to exposure to RF radiation. Each exposure session was preceded and followed by sham exposure sessions in which conditions were identical to exposure sessions except that the magnetron was not energized.

Results

Repeated exposures of the various animals to increasing power densities of radiation revealed levels where the animals' response rates were unaffected by the radiation, intermediate levels where response rates were marginally perturbed, and levels where responding eventually stopped. It was therefore possible to establish average power densities in each animal where disruption of behavior occurred. The medians of these various power densities were then used to calculate thresholds of disruption for the various animals and different experiments.

In addition, for the monkeys, colonic temperatures were simultaneously obtained that revealed the average temperatures of the animals corresponding to the power densities that produced response rate disruption. Temperatures were also obtained in some rat experiments but not simultaneously with the experimental sessions.

The thresholds for disruption of operant responding in rhesus monkeys in terms of power densities at frequencies of 225 MHz, 1.3 GHz, 2.45 GHz and 5.8 GHz were respectively 8.1, 57, 67 and 140 mW/cm^2. Similarly, the thresholds for the squirrel monkey at 2.45 and 5.62 GHz were 45 and 40 mW/cm^2, and for the rat at 1.28, 2.45 and 5.62 GHz were 10, 28, and 21 mW/cm^2. With few exceptions these power densities were all associated with an increase in colonic temperature of 1 °C or greater in the corresponding animals.

Discussion and Conclusions

Disruption of operant behavior by RF radiation is directly associated with increases in body temperature. Local increases in temperature may be the primary determinant of such disruption but we are unable to assess this aspect at the present time.

The various threshold power densities obtained in this series of studies provides sufficient data to generate a family of curves based on different characteristics of these animals and the frequency of the nonionizing radiation. Some of these curves illustrate linear and nonlinear relationships between nonionizing radiation and body mass, body surface area or diameter of the cranial cavity. All of these curves provide some basis, albeit tenuous, for predicting similar effects on humans of RF radiation.

References

de Lorge, J.O. The effects of microwave radiation on behavior and temperature in rhesus monkeys. In "Biological Effects of Electromagnetic Waves. Selected papers of the USNC/URSI annual meeting." Vol 1, U. S. Dept of Health, Education and Welfare, Public Health Service, FDA, Bureau of Radiological Health. Washington, DC: DHEW Publication No (FDA) 77-8010, 1976.

de Lorge, J.O. Operant behavior and colonic temperature of squirrel monkeys during microwave irradiation. Radio Science, 1979, 14(6S), 217-226.

de Lorge, J.O. Operant behavior and colonic temperature of rhesus monkeys, Macaca mulatta, exposed to microwaves at frequencies above and near whole-body resonance. Bioelectromagnetics, in press.

de Lorge, J.O. and Ezell, C.S. Observing-responses of rats exposed to 1.28- and 5.62-GHz microwaves. Bioelectromagnetics, 1980, 1, 183-198.

Holland, J.G. Technique for behavioral analysis of human observing. Science, 1957, 125, 348-350.

Knepton, J. and de Lorge, J. Effect of pulsed 5.62 GHz microwaves on squirrel monkeys (Saimiri sciureus) performing a repeated acquisition task. NAMRL-1291. Pensacola, FL 32508: Naval Aerospace Medical Research Laboratory, January, 1983.

Nelson, T.D. Behavioral effects of microwave irradition on squirrel monkeys (Saimiri sciureus) performance of a repeated acquisition task. NAMRL-1245. Pensacola, FL 32508: Naval Aerospace Medical Research Laboratory, March, 1978.

Reno, V.R. and de Lorge, J.O. A primate restraint device for use in microwave biological research. IEEE Transactions on Biomedical Engineering, 1977, BME-24, 201-203.

Sanza, J.N., and de Lorge, J. Fixed interval behavior of rats exposed to microwaves at low power densities. Radio Science, 1977, 12(6S), 273-277.

Acknowledgements and Disclaimers

Opinions and conclusions expressed in this report are those of the author and do not necessarily reflect the views or endorsement of the Navy Department. The animals used in this study were handled in accordance with the Principles of Laboratory Animal Care established by the committee on the Guide for Laboratory Animal Resources, National Academy of Science - National Research Council, DHEW Publ. No. (NIH) 80-23, 1980.

AD-P003 295

A Review and Evaluation of Research Concerning the Performance Effects of Nuclear Radiation

Ben B. Morgan, Jr., Richard B. Moe,
Frederick G. Freeman, and Christine B. Philput
Center for Applied Psychological Studies
Old Dominion University

Abstract

This report summarizes the findings of a review of studies of the effects of ionizing radiation on the task performances of primates. This review indicates that muscular and cognitive performance are more sensitive to radiation effects than are sensory-perceptual or psychomotor performances. Within performance domains greater performance decrements were found to occur with higher exposure levels. A need is indicated for future research to expand the range of exposure levels and the performance tasks employed.

Introduction

The modern world's increasing civilian and military usage of atomic energy and the concomitant increases in the possibility of human exposure to ionizing radiation, make it critically important that the behavioral effects of nuclear radiation be understood. In this regard, it appears that the relationships between radiation dosage and the nature and extent of the resulting physiological symptomatology are now fairly well understood. Recent reviews of the literature have provided a relatively clear picture of the symptomalogical sequelae that might be experienced by humans exposed to various levels of radiation (see Anno, Brode, & Washton-Brown, 1982; Baum, Anno, Young, & Withers, in preparation).

Very little is known, however, about the effects of radiation on human performance (cf. Morgan & Cruser, 1982). The available data suggest that exposures below 200 rad or 300 roentgens do not significantly impair the performances of patients who receive radiation therapy. However, this "suggestion" must be viewed with extreme caution because of the small number of studies on which it is based, and because of the fact that the experiments in question employed a very limited number of subjects, very low levels of irradiation, performance measures that have limited practical relevance, and subjects (radiation therapy patients) whose medical status was less than optimal. Thus, the relationships in humans between radiation exposure and the resulting cognitive and motor performance decrements remain largely speculative. This relative dearth of information is most significant for military planners because it fails to provide a basis for estimating the effects of battlefield irradiation on the combat effectiveness of military personnel. In order to circumvent this lack of knowledge, other sources of data must be sought as a basis for concepts and doctrine concerning nuclear warfare. One such source of data is the literature on the performance effects of radiation in animals. In spite of the importance of the animal data and the existance of a relatively long body of literature, relatively little attention has been devoted to summarizing the existing data as a basis for establishing the linkages between radiation and performance in humans. This paper reports the results of an effort to collect, categorize, and summarize the available primate data and to integrate the findings with knowledge of human performance research. Specifically, it summarizes the findings of a review and analysis of studies of the effects of ionizing radiation on the task performance of primates (see Morgan, Freeman, Winne, Moe, & Philput, 1984).

Method

A comprehensive review of the psychological, medical and radiobiological literature was conducted by executing both computer and hand searches of issues of Psychological Abstracts, Index Medicus, and available government documents between the years 1960 and 1983, inclusive. In addition, the NTIC, PSYCHINFO, MEDLINE, AND DTIC data bases were subjected to thorough computer searches. Liaison with the Radiation Branch of the USAF School of Aerospace Medicine (SAM), the Experimental Psychology Division of the Armed Forces Radiobiology Research Institute (AFRRI), and the Lovelace Foundation also provided important insight and valuable support for this review.

Approximately 250 citations and accompanying abstracts were reviewed for application to the specific areas of interest. Those considered germane were summarized and coded using a unique alpha-numeric descriptor that permitted catagorization along the following dimensions: type of subject; type of radiation exposure; total exposure; source (i.e. Cobalt-60, X-ray, etc.); type of task performed; training; length of data collection; symptomatology; and effects.

Measures

In order to document the effects of radiation in a way that would be relatable to human responsivity, three measures of individual responsivity to radiation were determined from the data; these were the number (or percentage of the sample) which displayed early transient incapacitation (ETI), permanent and complete incapacitation (PCI), or no effect (NE). In addition, two average or group performance measures were calculated; these were the amounts of early performance decrement (EPD) and recovery performance decrement (RPD).

Early transient incapacitation (ETI) was determined by examining individual plots of performance. In this report ETI is defined as a temporary, precipitous drop in performance within the first two hours after exposure. Researchers have usually defined ETI in more operational terms (e.g., incapacitation for a 5- to 10-minute period within the first half hour following radiation exposure). However, for the purposes of identifying the percentage of individuals that would not be able to perform immediately following exposure, the broader definition used here seems more appropriate. Generally, when ETI was exhibited, it occurred very soon (less than 1/2 hour) after exposure and only lasted for a few minutes. The performance of some animals, however, never recovered; these animals often died after a short time. A measure of Permanent and Complete Incapacitation (PCI) was used to characterize these subjects. Finally, there were some subjects that did not display any behavioral effects of radiation in the first two hours. These subjects were classified as exhibiting No Effect (NE). A percentage of the subjects falling into each of these three performance states was computed for each sample.

Early Performance Decrement (EPD) was defined as the mean level of performance during the initial 2-hour period following exposure (including periods of ETI). Data from tables and figures were used to calculate scores for each subject. These individual performance scores were then converted to mean percentage of baseline by dividing by the baseline performance level. EPD was then calculated by averaging across the individual subjects. When baseline measures were not reported, baseline performance was assumed to be 100%.

Recovery Performance Decrement (RPD) was defined as the mean level of performance following the period of EPD, from 2 hours post exposure until the end of each study. It was assumed that for subjects who were not totally incapacitated, performance following the first 2 hours post exposure would return to a level that reflected the longer-term impact of a radiation environment. To calculate RPD, tables or figures were used to identify the animals who were not totally incapacitated. Mean percentage of baseline performance was calculated using the same procedures discussed for the EPD measure. RPD was then computed by averaging across subjects.

Findings

Perhaps the clearest finding of this review is the fact that the performance effects of radiation in primates are not very straightforward. The results of a given study are determined by a variety of variables which seem to be manipulated differently in nearly every study. These variables include manipulation of independent variables along dimensions associated not only with the total exposure dosage but also with the source of the radiation, the gamma-to-neutron ratio of the source, the fractionation of dosage or use of multiple exposures, the dose rate of exposure, the part of the animals' bodies exposed to the radiation, and the dosimetry or measurement of the actual exposure. The interpretation of results is also complicated by variations in dependent variables along dimensions related to the type of task employed or type of performance measured, the amount of training provided the subject on the level of baseline established, and the specific criterion used to measure performance. Although all these factors must be considered in evaluating the results of previous studies, this report focuses on differences across different types of task performances by combining across (or ignoring) most of the other variables. Specifically, performance effects are summarized for four classes of performances; namely sensory-perceptual, cognitive, psychomotor, and muscular performances. These broad performance categories were chosen because they are often used in the classification of human performances (cf. Morgan & Cruser, 1982).

Sensory-Perceptual Performance

Nearly two-thirds of all studies of the radiation-stressed performances of monkeys involved the use of visual or auditory discrimination tasks. A perusal of these studies indicates that relatively little attention has been given to the effects of doses of less than 20 Grays (Gy 2000 rads). Nevertheless, results indicate that exposures of less than 20 Gy will produce relatively minor effects on simple discrimination performances; somewhat larger decrements were observed in time-critical performances, and these effects continued for several days post exposure. In the range of 20 to 50 Gy, sensory-perceptual performance effects are quite severe. Early performance decrements (EPD) range between 95 and 35% of pre-exposure levels, with an average of about 80% of baseline. Early transient incapacitation (ETI) is reported to occur in between 0 and 100% of the sample, with about 50% of the animals showing ETI in most studies. At the lower end of this range, there is some evidence of substantial recovery of performance. Above 50 Gy, all monkeys showed some effect, with ETI occurring in most. The levels of EPD ranged between about 88 and 32% of baseline, and in surviving animals, performance (RPD) recovered to about 75% of baseline. Thus, in sensory-perceputal performances there is a clear dose-response effect; however, the size of this effect seems to stabilize somewhat above 50Gy of radiation exposure.

Cognitive Performance

Very few studies have investigated the effects of radiation on what might be classified as cognitive performance in monkeys. Furthermore, all the studies in this area employed a match-to-sample or a delayed picture memory task with exposures of only 10 or 20 Gy. Thus, it is impossible to draw any conclusions concerning the dose-response effects of radiation on cognitive performance; a within-study comparison of 10 and 20 Gy showed essentially no difference in the performance effects of these two dosages. The major observation from these studies is that cognitive performance seems to be highly sensitive to radiation effects. Results from one study indicate EPD to be about 20% of pre-exposure baselines (with 25 to 37% ETI) followed by complete recovery of performance. Relative to other types of performance, these decrements are quite large. This suggests that cognitive tasks might be the most sensitive indicators of radiation effects. However, this suggestion must be viewed with caution and verified by additional research with a variety of cognitive tasks and exposure levels.

Psychomotor Performance

The assessment of complex psychomotor performance in monkeys has been accomplished almost exclusively with the Primate Equilibrium Platform (PEP). Radiation levels in the reported studies range from 3 to 27 Gy. In general, studies examining the effect of 15 Gy or less have reported only minimal levels of ETI, PCI, or EPD, with an apparently rapid recovery of performance to baseline levels. Early performance decrements in these studies were about 95% of baseline. Thus, it appears that, relative to cognitive performance, psychomotor performance is relatively insensitive to radiation effects, but about equal in sensitivity to sensory-perceptual performances at doses less than 15 Gy.

At doses of 25 and 27 Gy, more pronounced decrements in psychomotor performance are observed, with EPD reaching 48% of baseline in one study using 27 Gy. Thus, it appears that psychomotor performance is sensitive to radiation effects, but that relatively larger doses are required to produce reliable decrements.

Muscular Performance

Only the cognitive domain has been studied less than performances involving gross motor, muscular, or physical activities. The few studies in this area have employed the physical activity wheel or other tasks involving whole-body locomotion (e.g., maze traversing). All these studies have employed relatively high doses of radiation, with exposure levels ranging between 20 and 46 Gy. Taken together and compared with studies from other performance domains with similar exposure levels, these studies indicate that the performance of tasks involving muscular activity is more severely degraded than for sensory-perceptual or psychomotor tasks but not for cognitive tasks. Furthermore, muscular performances tends to be rather severely degraded by increases in radiation exposure levels there tends to be relatively little recovery of performance. For two studies involving 20 Gy exposure, the EPD averaged 55% of baseline, whereas in studies involving 40 and 46 Gy, the EPDs were 45 and 36%, and RPDs were 89 and 58%, respectively.

Summary and Conclusions

The results of this review indicate that very few studies have been conducted in the domain of cognitive performance--so few, in fact, that findings in this area must be viewed with extreme caution until confirmed by additional studies. Furthermore, it was found that at least two-thirds of all studies have been in the area of

sensory-perceptual performance. While additional research in this area may be warrented, it is clear that future research should focus on other performance domains, particularly the domains of cognitive and muscular performances.

This review also indicates that there has been little attempt to replicate findings with all conditions held constant. This makes it difficult to estimate the reliability of findings. Furthermore, there seems to have been no effort to determine the effects of a given exposure condition on a variety of task performances. Such research is clearly needed in order to enhance knowledge concerning the relative sensitivity of different task performances. It should also be noted that the available research is based on a very small sample of different performance tasks. In most cases, all research within a given performance domain has involved only one or two different tasks. While this narrow selection of tasks has great practical justification, it does limit the generalizability of findings. Questions concerning the generalizability of results are also raised by the finding that very large individual differences in radiation effects have been reported in nearly every study.

It was also found that except in the domain of sensory-perceptual performance, studies have covered a rather narrow range of radiation exposure levels. A majority of studies fall in the range of 10 to 30 Gy and few studies have involved radiation levels of less than 10 or more than 50 Gy. Thus, it is important that future research expand the range of exposure conditions used to test all except sensory-perceptual performances. Because the area of greatest concern with respect to the performance effects of radiation in humans is in the intermediate dose range, future animal studies should investigate more fully the effects of 3 to 10 Gy exposures.

Finally, it should be noted that within the limits of the available data, different classes of animal performance seem to be differentially sensitive to radiation effects. Specifically, sensory-perceptual and psychomotor performances seem to be least sensitive, whereas muscular and cognitive performances seem to be most affected by radiation. Furthermore, within a given performance domain there are apparently increasing effects of radiation as exposure levels are increased; however, this is hard to support in some cases (particularly in the cognitive domain) because of the restriction in range of conditions examined. Presumably, similar patterns of differential task sensitivity, dose-response effects, and broad individual differences in response should be expected in estimating the effect of radiation on human performances.

References

Anno, G. H., Brode, H. L., & Washton-Brown, R. (1982). Initial human response to nuclear radiation (PSR Note 477). Los Angeles, CA: Pacific-Sierra Research Corporation.

Baum, S., Anno, G. H., Young, R. W., & Withers, R. (in preparation). Symptomatology of the acute radiation effects in humans for doses from 0.5 to 30 Gy. Washington, DC: Defense Nuclear Agency.

Morgan, B. B., Jr., & Cruser, L. G. (1982). Performance effects of sickness and symptomatology: A review of the literature (Report No. 82-1-2). Norfolk, VA: Organization Research Group of Tidewater, Inc.

Morgan, B. B., Jr., Freeman, F. G., Winne, P. S., Moe, R. B., & Philput, C. B. (1984). A review of the performance effects of ionizing radiation in primates (ITR-84-28). Norfolk, VA: Center for Applied Psychological Studies, Old Dominion University.

AD-P003 296

Effects of Stress on Individual Productivity, Absenteeism, and Wellness

William H. Hendrix
Clemson University

Alex R. Rodriquez
OCHAMPUS

Abstract

The effects of job and life stress on productivity, absenteeism, and general wellness (number of cold and flu events per year) were investigated within a path analytic framework. Subjects were 207 individuals participating in a health promotion program in two DOD organizations in the Denver metropolitan area. The data indicated significant path relationships between life and job stress through job satisfaction and commitment to the factors of absenteeism and productivity. In addition, a series of stress related factors were significantly related to increased incidents of cold and flu episodes had per year.

Introduction

Health promotion and wellness programs are on the increase within organizations as individuals become more health conscious. Increased medical costs and absenteeism due to illness have encouraged organizations to actively seek ways to reduce these costs. Nationally, the cost associated with stress alone has been estimated to be approximately $75-90 billion annually (Ivancevich & Matteson, 1980). The U.S. Clearinghouse for Mental Health Information has estimated that U.S. industry has suffered a $17 billion annual decrease in productivity in the last few years due to stress related disorders (Yates, 1979). Stress has also been implicated as a factor which leads to job dissatisfaction, absenteeism and voluntary turnover (Ivancevich, Matteson & Preston, 1982) and to poor mental health (House, McMichael, Wells, Kaplan & Landerman, 1979; Kasl, 1973).

The present study focuses on the effect of life and job stress on productivity (merit performance appraisals) and absenteeism (sick leave rates). The major hypothesized model is provided in Figure 1. The model hypothesizes that life stress leads to job stress and job stress effects absenteeism both directly and indirectly through job satisfaction and organizational commitment. Also, job stress effects performance through its influence on job satisfaction and organizational commitment. In addition, it was hypothesized that the stress related factors of emotional exhaustion, job stress, anxiety, and anger/irritation were related to general wellbeing as reflected in the number of cold and flu episodes had over the last year. The number of cold and flu episodes provides a partial indication of the immune systems' effectiveness; the more episodes the lower the effectiveness. Research data has indicated that stress tends to increase the adrenal functions and decrease the immune systems' effectiveness in warding off diseases.

Method

Subjects

Subjects were 207 military and civilian employees from two DOD organizations located in the Denver metropolitan area who were participating in the OCHAMPUS Health Promotion Program. Approximately 33% were male and 67% were female. The individuals ranged in age from 20 to 70 with a mean age of approximately 43. Participation was on a voluntary basis and anonymity was ensured by using an individual identifier number, known only to the individual participant, for data collection.

Instrument

The Health Assessment Package was the principal survey used for data collection. The instrument consisted of approximately 308 attitudinal, behavioral, and background information items. They measured factors associated with the job environment, the extraorganizational (non-job) environment, individual characteristics, stress, job satisfaction, organizational commitment and health related factors. In addition, individual sick leave days taken over the past 6 months, and last merit performance appraisal scores were obtained for program participants.

Analysis

First, Path Analysis was performed to test the series of hypotheses depicted in Figure 1. Then correlational analysis was performed to test if the stress related factors of emotional exhaustion, job stress, anxiety, and anger/irritation were related to the number of cold and flu episodes had per year.

Results

Path Analysis

The first stage of this analysis involved regressing in a sequential process each factor in the model against all preceding factors. The beta weights obtained served as path coefficients and indicated the relative strength of the factors in the model.

For each regression analysis all factors not statistically significant ($p < .05$) were dropped out and then each regression analysis was re-run containing only the independent factors found to be significant ($p < .05$). This resulted in the revised path model depicted in Figure 2.

Correlational Analysis

The results of correlating the hypothesized stress related factors with the general wellness factor of number of cold and flu episodes had per year is provided in Table 1.

Discussion

Although many studies have focused on various effects of stress, few have used a multivariate approach. From researching the literature it appears that none have attempted to establish causality through the use of path analysis. It seems reasonable that the effects of stress on

absenteeism, and performance could have indirect effects by influencing one's job satisfaction and organizational commitment. This study investigated a series of hypotheses (depicted in Figure 1) which attempted to establish the path leading from life and job stress to absenteeism and performance.

A review of Figure 2 reveals that all of the hypotheses were supported except for one. The one not supported was that there was no direct relationship between job stress and organizational commitment (indicated by absence of a line in Figure 1). Each of the path coefficients listed in Figure 2 were significant at the .02 level or beyond. The data indicate that life stress in part directly increases one's stress on the job. One's job stress in turn directly influences one's job satisfaction and absenteeism rate. The higher the stress the more dissatisfied one is with their job and the more they tend to be absent from work. One's organizational commitment is directly influenced by job satisfaction and job stress. Therefore, job stress effects commitment directly (beta = -.20) and indirectly through its effect on job satisfaction. Performance, on the other hand, is only effected by stress indirectly through changes in commitment and job satisfaction. Therefore, both absenteeism and performance is effected by stress indirectly through changes in job satisfaction and organizational commitment, while absenteeism is in addition directly effected by job stress.

The data in Table 1 indicate that the job stress related factors of emotional exhaustion, job stress, anxiety, and anger/irritation, were all significantly ($p < .01$) related to number of cold and flu episodes had per year. That is, the higher each stress related factor the higher the incident of reported cold and flu events.

Taken together the data indicate that stress experienced on the job is in part influenced by one's life stress. In turn, job stress effects performance on the job, absenteeism rates, and one's general health as reflected in the increased occurence of flu and cold episodes for those experiencing stress.

Based on these results it appears that wellness and health promotion programs, which help individuals cope with stress and help identify job stress factors and reduce their adverse effects, should improve the well-being of employees while increasing organizational productivity.

References

House, J. S., McMichael, A. J., Wells, J. A., Kaplan, B. H., & Landerman, L. R. Occupational stress and health among factory workers. Journal of Health and Social Behavior, 1979, 20, 139-160.

Ivancevich, J. M., & Matteson, M. T. Stress and work: A managerial perspective. Glenview, IL.: Scott-Foresman, 1980.

Ivancevich, J. M., Matteson, M. T., & Preston, C. Occupational stress, Type A behavior, and physical wellbeing. Academy of Management Journal, 1982, 25(2), 373-391.

Kasl, S. V. Mental health and the work environment: An examination of the evidence. Journal of Occupational Medicine, 1973, 15, 509-518.

Yates, J. E. Managing stress. New York: Amacom, 1979.

Figure 1. Hypothesized Path Model

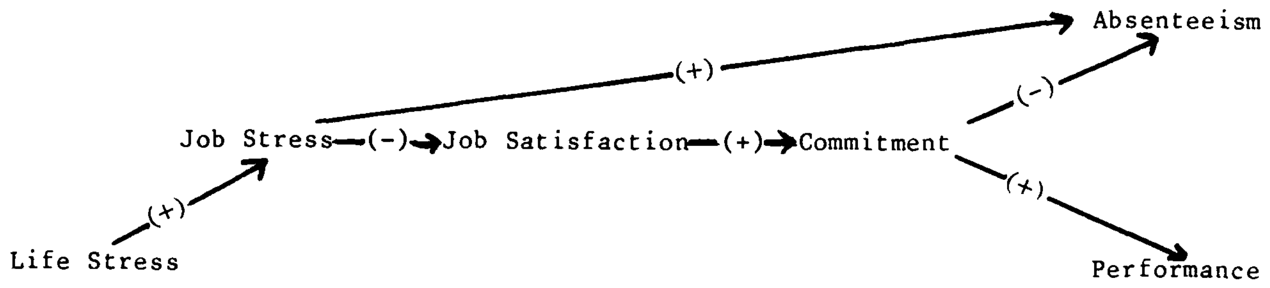


Figure 2. Revised Path Model

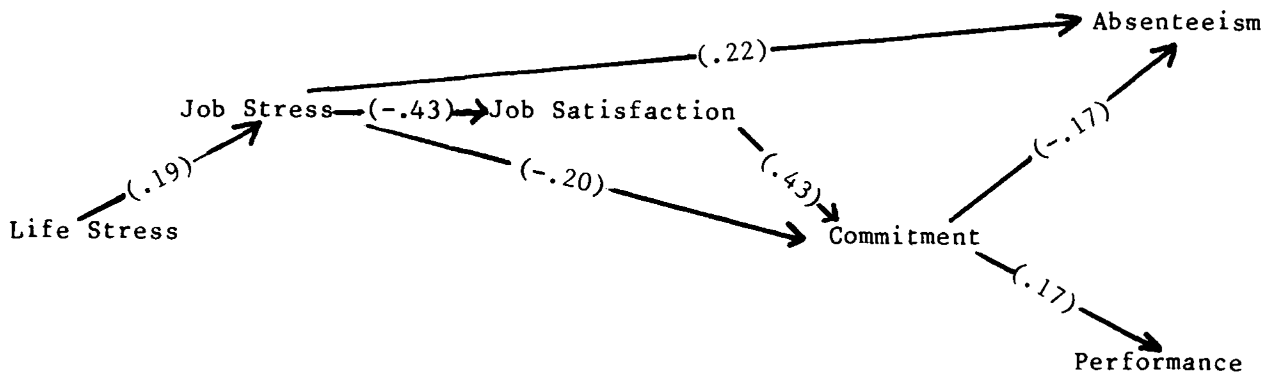


Table 1

	Emotional Exhaustion	Job Stress	Anxiety	Anger/ Irritation
No. Colds/Flu/Year	.19	.19	.23	.23

N = 207

STRESS AND MILITARY WOMEN: THE RELATIONSHIP OF JOB AND LIFE EXPERIENCES TO MENSTRUAL DISTRESS

Linda Torsani Fatkin
U.S. Army Human Engineering Laboratory, Aberdeen Proving Ground, Maryland

Traditional approaches to menstrual cycle research have concentrated on a disease-model framework. Emphasis has been on identifying the menstrual cycle as the independent variable that causes the events under investigation, or explains their variance. Parlee (1981) described this fundamental assumption as a deeply ingrained tenet of the medical-psychiatric approach. The context in which menstrual cycle variables, such as mood changes, occur is largely ignored, and their classification as normal or abnormal occurrences has appeared to be predetermined. Koeske (1981) emphasized that the behavior and moods of women are not ultimately and exclusively explainable as biological variable fluctuations, and that social and cognitive variables also need to be measured precisely.

There is a need for research detailing the contexts in which the menstrual cycle variables occur to provide the bases for a conceptual framework outlining areas of bio-social interaction. In accordance with this context-dependence framework, Sommer (1981) suggested as the selected sample for study, a high-risk group of women who might be more inclined towards experiencing menstrual distress because of their particular life history, life circumstance, or current situation.

The individual and circumstantial variables which influence the menstrual cycle are: (a) amount of situational stress, (b) age, (c) use of oral contraceptives and intrauterine devices (IUDs), (d) physical activity, (e) body composition, (f) subjective stress experience, and (g) femininity (e.g., Dan, Graham, & Beecher, 1980).

Two occupational groups were selected for the present study to represent traditional and nontraditional work environments that vary in degree of situational stress. It was predicted that women would report experiencing more menstrual distress if they: (a) worked in a nontraditional occupation; (b) experienced a high degree of occupational stress (defined by scores on the Job Related Strain Index, a greater proportion of men within their duty section, less job experience, or conflict between feminine values and job choice); (c) reported a high degree of life stress; and (d) reported a history of menstrual distress.

Method

Subjects

Subjects were 36 male and 25 female soldiers; 34 were military police (MP) and 27 were administrative specialists (ADMIN). Appropriate measures were taken to control for the following factors: (a) contraceptive use, (b) age range, and (c) percent body fat.

Materials

The General Background Questionnaire (GBQ) was devised for the present study and requested general background and job-related information.

The Life Experiences Survey (LES) was selected as a measurement of the stress of life changes. It is a 57-item self-report measure which allows subjects to rate separately the desirability and impact of events they have experienced during the past year (Sarason, Johnson, & Siegel, 1978).

The Inventory of Feminine Values (IFV) was used to measure attitudes toward the feminine role (Steinmann & Fox, 1979), and the Job Related Strain Index (JRS) was used as a measurement of job stress (Drew, 1982).

The occurrence of menstrual distress symptoms was measured by responses

on the Body Awareness Questionnaire (BAQ), modelled after the form used in the Wilcoxon, Schrader, and Sherif (1976) study. To avoid eliciting stereotypical responses, the specific intent of the study was not known to subjects and the BAQ instructions describe these items as experiences which men and women sometimes have. It allows subjects to describe the symptoms as they are experiencing them on the day the questionnaire is being answered, and is reported to have a high internal consistency and test-retest reliability (e.g., Koeske, 1981).

The Daily Events Rating Scale (DERS) is a short form attached to the BAQ. The form was devised for the present study to allow subjects in each occupation to assign a subjective stress rating to the events which occurred that day.

Procedure

The men and women attended orientation meetings in mixed groups of five, and direct reference to menstruation was avoided by referring to the study in the context of general health issues. Subjects who agreed to participate for 35 consecutive days signed the consent forms, and completed the four initial questionnaires: the GBQ, the LES, the JRS, and the IFV. They were then given an envelope containing the daily questionnaires (the BAQ and the DERS). Subjects were instructed to complete the daily questionnaires at the end of their work day and return the sealed envelope to a central (and convenient) location before leaving for work the following morning. They obtained the next envelope at the same location upon their return from duty each day.

Individual End-of-Study Interviews included specific questions which requested the subjects' comments about the study, their perception of the validity of their test scores, and other job-related questions. In addition to the standard questions, women were asked about the regularity of their cycles before entrance into military service, specifically any history of menstrual distress.

Results

Analysis

The predictor variables included in the multiple regression analysis were: (a) job stress (measured by the JRS), (b) feminine values (IFV), (c) life stress (LES, negative score), (d) subjective stress (Overall DERS score), (e) job, (f) job experience, and (g) history of menstrual distress. The outcome variable was the menstrual distress score, which was derived from the mean score of the daily BAQ's. All significance levels for F ratios were set at $p<.05$.

Traditional vs. nontraditional occupations. Women working in the nontraditional occupation of Military Police (MP) reported significantly more menstrual distress than women working in the traditional administrative positions (ADMIN).

Correlation matrix. A significant relationship was found between life stress scores and menstrual distress scores ($r=-.5619$). Those who experienced more negative life stess, also reported more menstrual distress.

There was a significant correlation between job and reports of menstrual distress ($r=-.4295$). Women in the nontraditional job (MP's) were more likely to report distress than those who worked in the traditional occupation.

Level of job experience and reports of menstrual distress were also significantly correlated ($r=-.4762$). Women who were in the lower job experience level (7 months or less) reported more menstrual distress than those with more job experience (8 months or more).

Job stress and job experience level were correlated in a similar manner ($r=-.4384$). Women in the lower experience level reported more job stress

than women who had worked in their jobs for longer periods of time.

Type of job and level of experience were also correlated (r=.5192). Thus, when dividing the women according to job and experience level, the MP's working at low levels of experience outnumbered the Admin women. The reverse was true for the higher experience level.

Regression analyses. In order to select the minimum number of variables necessary to account for the maximum amount of variance in the total set, a forward regression solution was used. Life stress (accounting for 32% of the variance), job experience, subjective stress, job, feminine values, and job stress were the six predictors that accounted for a significant amount of the variance on menstrual distress. A multiple regression analysis was then performed on these dominant predictors, yeilding a multiple $\underline{R}$ of 0.6956, $\underline{F}(6,18) = 2.81$, $\underline{p}<.05$.

Post-hoc Comparisons

Sex differences. Males were included as a comparison sample because their behavior has been the implicit standard against which feminine fluctuations have been compared. The outcome variable is discussed as "health distress" because the term "menstrual distress" is applicable only to females.

MP females reported the highest amount of health distress and job stress. However, the life stress reports seem to vary according to job rather than sex. The MP men and women rated their life experiences as more stressful than the ADMIN group.

Females gave their daily events higher (more positive) ratings than the males, as indicated by their subjective scores. Males had perceived a larger proportion of the 35-day rating period as being negative, which resulted in the lower mean scores.

Discussion

Methodological Issues

The present study incorporated many of the methodological improvements emphasized for menstrual cycle research by other investigators (e.g., Dan et al., 1980). A non-clinical population was used, along with an initial cover story that avoided direct reference to menstruation. These factors were used to avoid the problem of demand effect and volunteer bias. Questionnaires were collected on a daily basis to avoid reliance on retrospective measures of body symptoms and to prevent referrals to previous reports completed by the men and women. Undertaking the study of menstrual distress within a context-dependence framework, provided for the analysis of factors which have an affect on the degree of menstrual distress for a specific subgroup of our population, rather than assuming ubiquitous effects for all women.

Job and Menstrual Distress

The increased reports of menstrual distress in the MP women may be due to a greater sensitivity and awareness to environmental occurrences. Kanter (1977) emphasized that women in male-dominated jobs frequently share experiences--the organizational discomfort of being pioneers, the feelings of isolation, the lack of support from male colleagues, loneliness, and sexual discrimination. Although these experiences may not lead to a dissatisfaction with the actual duties on the job, women may instead be more attentive to these external pressures and to changes in their body states.

The changing shift routine was identified as a primary stressor for the MP women. The MP's in the present study worked under a three-day rotating shift work schedule consisting of three "day shifts" (0700-1500 hours), three "swing shifts" (1500-2300 hours), three "midnight shifts" (2300-0700 hours), and three days off. The scheduling of shift work profoundly affects

psychophysiological processes. Luce (1971) reported that rotating shifts put a person in a state of internal desynchrony; some individuals have shown an adjustment to night shift within 5 days, while others have not adjusted in 21 days. The psychological strains and physiological changes that can occur during rotating shift schedules could lead to an increased vulnerability for the MP's so that they become more susceptible to the experience of menstrual distress.

The MP males had similar reactions to the 3-day rotating schedules, including complaints about the repeated disturbances in their sleeping and eating patterns. This result was not surprising as Unger (1975) has reported that women are not alone in having biological cycles that are vulnerable to shift changes. An increasing number of biological parameters have been shown to vary in both sexes with time of day and over weeks and months.

Parlee (1981) emphasized that male performance is not irrelevant to the study of behavioral changes associated with the menstrual cycle because it is important to the study of "rhythmic changes in human behavior... a more useful concept in a general psychological theory". Research directed toward "human cycles" is not intended to deny the reality of the menstrual cycle or the various symptoms and changes in behavior associated with it, but serves to place the menstrual cycle in its proper perspective along with many other biological cycles that affect the lives of both males and females (Unger, 1975).

Menstrual Distress and Degree of Occupational Stress

Level of job experience was considered to be a contributing factor to the perception of the occupational stress and reports of menstrual distress. Individuals with less experience might experience uncertainty about control over job-related events; this uncertainty has significant effects on health (Suls & Mullen, 1981). Those individuals with more skills and broader experience act more confident and seem to cope adaptively with a current stressor (Rabkin & Struening, 1976).

In the present study, the assumption that a reciprocal relationship exists between amount of job experience and reports of menstrual distress and job stress is premature. Because of the significant correlation found between job and job experience, the possible effect of job experience on menstrual distress and on job stress scores is confounded by the influence of the unequal proportions of MP subjects and ADMIN subjects in each experience level. It is probable that the higher menstrual distress scores and job stress scores in the low experience group are related to the higher proportion of MP's in that group, and not the amount of job experience. This issue still needs to be investigated to determine if the larger proportion of MP's in the low experience group could be corrected through more stringent experimental control, or if the small amount of time-on-the-job is in fact a by-product of the job itself.

Life Stress and Menstrual Distress

While scores from the LES were found to be significantly correlated with the menstrual distress scores, other social-environmental factors must be acknowledged when assessing the relationship between the life stress scores and reports of menstrual distress. For example, subjects in the MP group have a life-style of rapid changes, requiring many adjustments and very little time to adjust effectively. They are expected to deal professionally with unexpected events; they are to maintain their authoritative image; they are to fulfill complex and varied duty assignments; and they must cope with the 3-day rotating shift-work schedule.

The inconsistency and lack of personal control in the MP's environmental

and occupational life-style may have the greatest impact upon the experience of menstrual distress. The implication that social-environmental factors are indeed significant contributors to menstrual distress for military policewomen should be explored further to test the generalizability of this theoretical issue. To what other nontraditional occupations can this effect be generalized?

A practical application of these research findings is that the identification of organizational stress factors that affect the emotional and physiological health of specific personnel subgroups may be facilitated. Effective organizational changes are made through a careful consideration of the employees' well-being as well as an assessment of organizational needs. These changes may not only prevent possible long-term deleterious effects on health, but may also promote the optimum utilization of male and female personnel.

References

Dan, A.J., Graham, E.A., & Beecher, C.P. (Eds.). (1980). The menstrual cycle: A synthesis of interdisciplinary research (Vol. 1). New York: Springer.

Drew, K.E. (1982). Menstrual symptoms and occupational strain in working women. Unpublished masters thesis, University of Illinois.

Kanter, R.M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. American Journal of Sociology, 82, 965-990.

Koeske, R.D. (1981). Theoretical and conceptual complexities in the design and analysis of menstrual cycle research. In P. Kommenich, M. McSweeney, J.A. Noack, & N. Elder (Eds.), The menstrual cycle: Research and implications for women's health (Vol. 2). New York: Springer.

Luce, G.G. (1971). Body time: Physiological rhythms and social stress. NY: Pantheon Books.

Parlee, M.B. (1981). Gaps in behavioral research on the menstrual cycle. In P. Komnenich, M. McSweeney, J.A. Noack, & N. Elder (Eds.), The menstrual cycle: Research and implications for women's health (Vol. 2). New York: Springer.

Rabkin, J.G., & Struening, E.L. (1976). Life events, stress, and illness. Science, 194, 1013-1020.

Sarason, I.G., Johnson, J.H., & Siegel, J.M. (1978). Assessing the impact of life changes: Development of the Life Experiences Survey. Journal of Consulting and Clinical Psychology, 46, 932-946.

Sommer, B. (1981). Menstrual cycle research: Yesterday, today, and tomorrow. In P. Kommenich, M. McSweeney, J.A. Noack, & N. Elder (Eds.), The menstrual cycle: Research and implications for women's health (Vol. 2). New York: Springer.

Steinmann, A.G., & Fox, D.J. (1979). Manual series for the interpretation of the Maferr Inventory of Feminine Values. New York: Maferr Foundation.

Suls, J., & Mullen, B. (1981). Life events, perceived control and illness: The role of uncertainty. Journal of Human Stress, 7, 30-34.

Unger, R.K. (1975). The case for human cycles. In R.K. Unger, & F.L. Denmark (Eds.), Woman: Dependent or Independent Variable? New York: Psychological Dimensions.

Wilcoxon, L.A., Schrader, S.L., Sherif, C.W. (1976). Daily self-reports on activities, life events, moods, and somatic changes during the menstrual cycle. Psychosomatic Medicine, 38, 399-417.

AD-P003 298

A Review of Open Literature and Recent DOD-Funded Research on Stress in Organizations

James T. Lester
Office of Naval Research

Abstract

One hundred and forty articles from refereed journals, in abstract form, were reviewed. The results of this review are outlined under the headings: sources of organizational stress, outcomes of organizational stress, moderating variables, and miscellaneous other studies. Recent Department of Defense (DOD)-funded efforts in this area are surveyed.

Introduction

Background

The study of stress in the work place appears to have gone through a development much like that of the field of industrial psychology in general, starting with an interest in blue-collar level employees, and moving toward attention to higher and higher levels of management and to organizational issues. The earliest work was concerned with specific objective stressors, of the kind likely to be factors in the work day of the blue-collar worker e.g., noise, temperature extremes, lighting inadequacies, and of course work-load. These factors were considered potential degraders of work-task performance. As Selye's conception of physiological stress caught on, there was growing research interest in physiological measures as dependent variables. The physiological outcomes tied in with a concern for health and well-being rather than with performance and productivity, and task performance as an output variable moved somewhat into the background.

The concept of stressors expanded to include psychosocial aspects of the victim's world, and began to take into account not just the objective, physical aspects of the environment but also, and perhaps more importantly, the subject's interpretations of his or her environment, especially the social environment and the important relationships in which he or she was involved. On the outcome or dependent variable side, basic physiological measures were joined by emotions and generalized emotional reactions to the work situation and to specific aspects of it. The most recent result of this development is the focus of attention on the subject of organizational stress, or a concern for those psychosocial factors within an organization which affect individual well-being.

Recently I carried out a review of the scientific literature bearing on the topic of organizational stress. This presentation includes an outline of this review and its findings, and an overview of recent and current work funded by the DOD on this topic.

Published Reports on Organizational Stress, 1974-1982

Sources of organizational stress

The majority of studies found seem to focus on the question, what things can serve as sources of stress in organizations? A great variety of possible stressors have been considered, many of them specific to certain occupations, but clearly one large sub-cluster concerns the topic of role-related stress. Ambiguity of role, conflict of roles, over-load of roles, these are stressors to which employees in almost any organization might be exposed, and illustrations of the effects of role-stress on well-being have been produced for manufacturing managers, R&D professionals, VA hospital staff, supermarket managers, mining employees, chief executives of public corporations, sailors, athletes and policemen, to mention only some.

Other findings on sources of stress do not cluster so well. A partial listing of suggested evils would include personnel practices, a tall organizational-hierarchy structure, a perception of administrative irrationality above one, a lack of job mobility, conflict over policy, geographical relocation, and rapid changes of various kinds.

One way to summarize these various studies is to say that a wide variety of things have been labeled as stressful, by either the subjects or the investigators. My own conceptual filtering system tells me that a general factor exists among these many things, which might be called the perceived feasibility of meeting expectations. The implicit outcome for individuals experiencing role conflict, ambiguity, or overload, and for individuals hampered in their efforts to fulfill expectations or reach goals will often be frustration or anxiety, and it seems to me these are the psychological concepts that pull back together the small explosion of answers so far found to the question of what it is in organizations that cause stress.

Outcomes of organizational stress

What sort of outcomes do we see through the eyes of researchers? First and foremost one sees expressions of tension or dissatisfaction. In fact, it is usually such expressions that lead investigators to label a condition as "stressful." Beyond this, there is a domain of physiological variables that show responses to stressors, and there is the whole domain of health indicators. There was no explicit attempt in this review to tap these domains although their literature is large. Other more psychological variables which have been reported to show observed relationships to stress level include: the basis on which decisions are made; degree of social comparison going on within a group; perception of relations within a group; leader's appraisal of subordinates' performance; employee's satisfaction with co-workers; cognitive fatigue; absenteeism; turnover; increased negative attitude to role-senders; over-assertion of authority; emotional withdrawal; and, de-valued self-conception. This review of reported stress-related outcomes left me with a sense of many disconnected bits and pieces, and of a crying need for integration via theory. We need someone to persuade us which effects are primary and which secondary, and to model the mechanisms by which both come about. There is something interesting here, but it isn't even clear (to me) whether or not the concept of stress is the best way to pull it all together.

Variables moderating stress in organizations

Almost all studies of stress report wide-ranging individual differences; that is, if one holds the source of stress constant or limits it to one factor, then both the degree and kind of reactions to the stressor vary considerably among persons. Or, if one studies one particular stress reaction and seeks its antecedent causes, one usually finds a spectrum of antecedents among individuals. One way to try to bring some order into such observations is to go beyond the input--output, or stimulus--response paradigm for studying stress, and give consideration to possible moderators of the input--output relationship. For example, one author obtained questionnaire results from a large group of school administrators, and found that female school administrators reported less occupational stress than did males. Social support has frequently been proposed and observed as a moderator. So have various personal characteristics, the most frequently mentioned of which is the "locus of control" personality dimension, and in the health-related studies, the "Type A" personality dimension. Even traits valued by the culture may play an undesirable role from the point of view of stress; one author reports finding high idealism and a single-minded sense of life-purpose more characteristic of people who were described as cases of occupational burn-out. The set of references under review included at least 37 different potential moderator variables, almost all of which have demonstrated the ability to moderate the relationship between at least some stressor and some outcome.

Miscellaneous other studies

Coping. One subject that might have been subsumed under moderator variables is that of coping: some investigators, but not many, have tried to study what people do about the stressors they find acting on them, on the assumption that one's coping method may influence the effects a stressor will have. It is not easy to provide researchable definitions of "coping", and researchers appear to have relied heavily on simply asking people how they typically deal with tension. Lists of coping actions obtained this way are then clustered into classes of actions, but no generally applicable taxonomy of coping strategies has been achieved. One theory groups coping behaviors into two large clusters: task-centered vs. emotion-centered; another into three: hypervigilance, vigilance, and defensive avoidance. Other strategies identified include: taking direct action; rationalizing or avoiding the stressor; accepting the situation as it is; seeking social support; decreasing affiliative behavior; turning to narcotizing agents; and, relying on professional help.

Stress-management interventions. Many stress-management ideas are proposed; almost none have received a clear empirical evaluation. Relaxation techniques are important elements in most interventions. Most programs concentrate on helping the individual to monitor and control physiological responses; on identifying trouble-causing responses (from teeth-grinding to apathy or anger) and changing them, usually by behavioral modification or biofeedback techniques; and on altering the person's ways of looking at himself or herself and the surrounding situation. The best programs neatly bypass some of the problems with the knowledge base by urging each person to find his or her own individual stressors and stress-related responses, and encourage each person to find what works for him or her as coping and

stress-management strategies. There have been very few laboratory or quasi-experimental studies of stress-management techniques.

Some concluding comments

The concept of stress has so far proved productive of studies and elusive of definition. It serves a useful function in everyday language, by giving us an objective-sounding, "scientifically" credible explanation for feelings of tension and medical symptoms of unclear origin. The concept gives employees in organizations a way to focus management attention on conditions that otherwise might be dismissed. But the current "scientific" credibility is like paper money - it needs something of real value to back it up. The melange of interesting studies on organizational stress has so far not been made to coalesce within a persuasive theory.

Recent DOD-Funded Work on Organizational Stress

(The presentation included a survey of recent DOD-funded work, but this was not available at the time the Proceedings went to press. Interested readers should contact the author - J. Lester, ONR-442OE, 800 N. Quincy St., Arlington, VA 22217 - for a written summary of this portion of the presentation.)

Aftereffects Associated with One or Two Stressors Across Conditions of Complete, Partial, or No Control

Captain Daniel L. Collins, A. Baum, and J. E. Singer
Uniformed Services University of the Health Sciences

Abstract

This study examined human aftereffects (i.e., psychological, behavioral, and physiological) associated with one or two stressors (i.e., noise, noise and/or a strobe light) over three levels of control (i.e., no, partial, or complete) where control could be repeatedly exerted to terminate the stressor(s). The new concept of partial control, rather than complete or no control, is more analogous to the many daily interactions humans face. Greater psychological, behavioral, and physiological aftereffects were associated with two stressors or the uncontrollable conditions; partial or complete control conditions were analogous to the comparison (no stressor/no control) group. Urinary catecholamines (i.e., epinephrine and norepinephrine) paralleled these findings showing greater arousal when control was absent than when control was available. Men had significantly greater catecholamine levels than did women.

Introduction

Background

In the past decade research has focused on the aftereffects of stress. Aftereffects are traditionally thought of as those behavioral anomalies that are found immediately after the stressor is terminated, such as a decreased tolerance for frustration, diminished proofreading performance, and less ability to deal with response competition (Glass & Singer, 1972). The term "stressor" is used to refer to stimulation that represents an adaptive threat or potential adaptive threat to the organism. The present study was concerned with two major issues; behavioral and physiological aftereffects associated with complete, partial, or no control over terminating the stressor(s) and aftereffects associated with exposure to one or two stressors.

Method

Two different stressors were used in this study. The first was a loud noise (95 dba) delivered over headphones. The second was a flashing strobe light placed at eye level approximately two feet away from the subject's head. Crossed with stressor conditions were varying conditions of complete, partial, or no control. For each combination of stressors, there was a control condition and a no control condition (e.g., could or could not terminate the stressors). Partial control referred to a condition in which noise and light were presented but subjects could only terminate the noise. Complete control referred to conditions in which one or two stressors were presented and subjects could terminate the stressor(s). No control referred to conditions in which one or two stressors were presented but could not be terminated.

Procedures

All subjects were recruited and randomly assigned to conditions and were paid \$10.00 for their participation. Informed consent was obtained in accordance with the University's guidelines for human subjects. Two urine samples were obtained, one before and one after the session; these were preserved for later assay. Subjects provided demographic information and were instructed in the use of the equipment.

Next, the subjects solved aloud a number of machine paced mental arithmetic tasks, while being exposed to manipulations of control and exposure to stressor(s). Each presentation of a stressor was unpredictable (e.g., no warning of its onset) and each was presented 18 times for 8 seconds at predetermined random intervals throughout the 21 minute session. Subjects either exerted control or attempted to do so, by pressing calculator type keys in their efforts to terminate the stressor(s). The subjects in the control conditions were informed of the sequence necessary to terminate the stressor(s). Subjects in the no control conditions were told some sequence of numbers would terminate the stressor(s) when in fact no solution existed.

Immediately following the arithmetic problem presentation, a five minute proofreading task was administered, followed by a persistence task and a visual perception task. The proofreading task involved identifying and circling misspellings, grammatical mistakes, incorrect punctuation, and typographical errors. The timed decision-making task required determination of which two shapes was the longer. The object of the persistence task was to predict the next letter that would logically follow a previous sequence of letters. The subjects were presented with up to 16 letters one at a time and could terminate this solvable procedure whenever they chose. The subjects proofreading performance, time to reach a decision on the decision-making task, and the time spent on the persistence task, were used to detect the presence of aftereffects. Additionally, psychoendocrine measures of poststressor arousal were measured using a double urinary void, through which change scores were derived. A final questionnaire, which included manipulation checks, was completed prior to debriefing the subjects.

Results

It was predicted that greater aftereffects would occur as the number of stressors increased and control decreased. The design of this study required data to be analyzed using planned comparions and are presented as means with higher values reflecting more of the given dimension. All catecholamine data are reported as changes in excretion levels (in ng/ml).

Manipulation Checks

Two manipulation checks were used in this study. The first manipulation check asked the subjects to indicate the number of stressors (i.e., 0, 1, or 2) to which they had been exposed during presentation of the arithmetic problems. All subjects correctly identified the number of stressors. The second manipulation check asked the subjects to indicate the amount of control they felt they had over the stressor(s). Perceived control was associated with the manipulations of control.

Replication of aftereffect findings

Previous research (Cohen, 1980; Glass & Singer, 1972) has demonstrated aftereffects following exposure to stressors by asking subjects to work on proofreading and persistence tasks. Using the same arithmetic problems and poofreading task as Glass and Singer (1972), similar aftereffects were found in the present study. Subjects who were not exposed to any stressor found more errors ($\bar{x}$ = 64.9%) in the passage that they proofread than did subjects exposed to one or more stressor(s) ($\bar{x}$ = 52.7; t (64) = 2.45 $p < .017$).

Glass and Singer (1972) also found that having perceived control over a stressor reduced aftereffects of exposure to the stressor. The present study provided subjects with actual control which was repeatedly exerted, and found the same pattern of aftereffects reported by Glass and Singer (1972). Comparison of subjects' proofreading performance in the no control ($\bar{x}$ = 48.95%) and complete control conditions ($\bar{x}$ = 59.55%) revealed significant aftereffect differences, t (64) = 2.45, $p < 0.17$. Having control, even when the control required work or effort in order to be maintained, was associated with better performance than was not having any control over the stressor. This pattern was also revealed on the persistence task. Subjects in the no control conditions persisted less ($\bar{x}$ = 85.45 seconds) than did those subjects in the conditions of complete control ($\bar{x}$ = 164.55 seconds), t (64) = 5.37, $p < .001$.

The present study sought to expand research on aftereffects associated with exposure to stress in several ways. First, the effects of exposure to more than one stressor were considered. As predicted, in the no control conditions, subjects' proofreading performance was better when one stressor was present ($\bar{x}$ = 56.0%) than when two were used ($\bar{x}$ = 41.9%) t (64) = 2.49, $p <$ .015. Similarly, persistence decreased as the number of stressors increased ($\bar{x}$ = 164.5 seconds vs. 85.5 seconds), t (64) = 5.37, $p < .001$.

When control was provided over both stressors, the effect was less dramatic. Exposure to one stressor-with control was associated with better proofreading performance ($\bar{x}$ = 65.5%) than was exposure to two stressors-with control of both ($\bar{x}$ = 53.6%) but this was only marginally significant t (64) = 1.82, $p < .07$. There were no significant differences on the persistence task between the two stressors with control ($\bar{x}$ = 152.7 seconds) or one stressor with control ($\bar{x}$ = 176.4 seconds) conditions, t (64) = 0.58, $p < .57$.

Thus far we have shown that the effects of perceived and actual control in laboratory studies of stress are comparable and that exposure to two stressors has more negative aftereffects than does exposure to one stressor, primarily when control is not available. As noted earlier, research on aftereffects has been restricted to conditions of either no control or complete control. This study additionally examined the aftereffects of exposure to stress under partial control. Significant aftereffect differences were noted between the subjects in the two stressor-partial control and the two stressor-no control conditions. Subjects in the two stressor-no control condition performed more poorly on the proofreading task ($\bar{x}$ = 46.7% vs. $\bar{x}$ = 41.9%) t (64) = 9.72, $p < .001$, and their persistence was less ($\bar{x}$ = 120.2 seconds vs. $\bar{x}$ = 67.4 seconds) t (64) = 4.15, $p < .001$ than subjects in the partial control condition. Similarly, the subjects' proofreading performance in the two stressor-no control condition was significantly lower ($\bar{x}$ = 41.9%) than the subjects in the two stressor condition with control ($\bar{x}$ = 53.6%), t (64) = 11.15, $p < .001$. This suggests that partial and complete control are

associated with fewer or less severe aftereffects than the multistressor-no control condition. Analyses between the partial control and the two stressor-control conditions revealed no significant aftereffect differences on any of the dependent variables.

The present research also examined psychoendocrine measures of arousal during the session. The catecholamine levels showed a main effect for control and for gender. The epinephrine (i.e., change scores) were significantly greater in the no control conditions than in the complete control conditions $\underline{t}$ (48) = 2.49, $\underline{p}<.015$. Analyses of these changes between the partial control and the no control conditions indicated greater physiological arousal in the latter conditions, $\underline{t}$ (28) = 2.23, $\underline{p}<.03$. Epinephrine levels reflected significantly greater increases for men ($\bar{x}$ = 7.12 ng/ml) than for women ($\bar{x}$ = 2.15 ng/ml) $\underline{t}$ (28) = 23.07, $\underline{p}<.001$, in the no control conditions. When subjects had control over terminating the stressor(s), the epinephrine levels between men and women were not significantly different $\underline{t}$ (18) = .64, $\underline{p}>.05$. Collapsing across conditions, men's epineprhine levels were greater ($\bar{x}$ = 4.6 ng/ml) than were the women's ($\bar{x}$ = 2.6 ng/ml), $\underline{t}$ (48) = 14.49 $\underline{p}<.001$. These analyses revealed a main effect for conditions of control as well as for gender. Norepinephrine change scores reflected similar trends.

Discussion

This study suggests that when control is available, exposure to environmental stressors results in few, if any, aftereffects. The post-stressor effects of exposure to none, one, or two stressors were measured across conditions of complete, partial, or no control. Partial or total control was assocaited with fewer poststressor decrements than were the no control conditions. Further, multistressor conditions were associated with greater aftereffects than were single stressor conditions. These findings have a number of implications for understanding both environmental stressors and their impacts on behavior, performance, and safety. The additional work of exercising control, even if it is only partial control, appears to have ameliorative effects that outweight the negative aspects of increased responsibility.

The notion of uncontrollable stress resulting in a sense of helplessness has been explored in various research settings (Seligman, 1975). Seligman, Maier, and Geer (1968) have proposed a theory of Learned Helplessness. This theory plays a central role in explaining our subjects behavioral differences and basically states that when the outcomes are independent of the inputs, the organism learns to be helpless. Thus, incentives for initiating actions aimed at avoiding the aversive event are absent. Conversely, when the subjects are able to control these aversive stimuli, they will be facilitated in performing their escape and avoidance behaviors, thereby avoiding the addiitional anxiety produced by the feelings of learned helplessness. In both cases, the subjects' behavior can be viewed as adaptive.

Interestingly, it appears that it is not the stressor(s) per se, that produces discontinuity of performance, but rather the individual's perception of control in the context of stressful experience, which can later effect behavioral functioning. To understand the psychological processes behind these differential behaviors has implications for all segments of the military and society. While the present study has expanded the current knowledge on aftereffects to include two stressors, and a partial control condition, more laboratory and field research is needed before we fully understand the impact

that aftereffects have upon our behaviors, health or in dubious acts against society.

References

Cohen, S. (1980). Aftereffects of stress on human performance and social behavior: A review of research and theory. Psychology Bulletin, 88, 82-108.

Glass, D. C., & Singer, J. E. (1972). Experiments on noise and social stressors. New York: Academic Press.

Seligman, M. E. P. (1975). Helplessness: On depression, development and death. San Francisco: W. H. Freeman & Co.

Seligman, M. E. P., Maier, S. F., & Geer, J. (1968). The alleviation of learned helplessness in the dog. Journal of Abnormal and Social Psychology, 73, 256-262.

AD-P003 300

PSYCHOLOGICAL AND BIOCHEMICAL EFFECTS OF A STRESS MANAGEMENT PROGRAM

T.Cayton[1],J.Patterson[2],W.Pierson[1],and G.Troxler[2]
[1] Wilford Hall USAF Medical Center and
[2] School of Aerospace Medicine

Abstract---15 female and 11 male volunteers were administered a battery of tests measuring 19 biochemical, psychological, and behavioral risk factors associated with the effects of stress pre- and post-participation in a 6 week stress management program. Significant pre/post differences at a probability level of at least less then .05 were found on 8 measures. Data were then grouped according to level of cardiovascular risk as determined by the total cholesterol to HDL ratio above or below 6. Results showed no significant differences in percent change by risk group except in the areas of HDL and ratios in which the high risk group had greater therapeutic changes.

For 1982, the Medline data base listed 959 references under the subject heading of psychological stress. Stress has been shown to have many detrimental health effects. It has been linked to the exacerbation of many medical problems including: cardiac arrhythmias (Lynch, Paskewitz, Gimbel, & Thomas, 1977), heart disease (Glass, 1977), diabetes melitus (Tarnow & Silverman, 1981), sudden death (Engel, 1978), immune system suppression (Locke, 1982), cancer (Sklar & Anisman, 1981), depression, hypertension, anxiety, and others (Borysenko, 1982; Dohrenwend & Dohrenwend, 1974; McQueen & Siegrist, 1982).

Many stress programs are being offered to the public. Few of these programs, however, have systematically evaluated their impact on psychometric, behavioral, and biochemical stress measures. Those investigating the effects of stress management have primarily used measures of trait anxiety only (Charlesworth, Murphy, & Beutler, 1981) or EMG, finger temperature, and state/trait anxiety (Allen & Blanchard, 1980), or were ineffective in reducing cholesterol or blood pressure levels (Jenni, & Wollersheim, 1979).

The present study is the first of a series of investigations to assess the effects of stress management on selected psychological and biochemical risk factors for cardiovacular disease. We wanted to sample emotional, behavioral, physical, and biochemical measures which might be sensitive to the changes of an effective stress management program.

The State/Trait Personality Invertory (STPI, Spielberger, et.al., 1979) and the Zung Depression Scale (Zung, 1965) were selected because of their usefulness and brevity. "Your Lifestyle Profile", a list of 35 questions covering behavioral risk factors such as smoking, activity level, etc. and the Jenkins Activity Survey (JAS), a psychometric measure of Type-A behavior, were included to assess changes in behavioral risk factors. In the realm of physical measures, heart rate (HR) and

blood pressure (BP) were included. In addition, we asked the subjects to indicate if they had been recently (one month) bothered by any of a list of approximately 100 physical symptoms.

With the exception of HR and BP, the above measures are easily susceptible to demand characteristics. Therefore, biochemical data associated with the stress response, cortisol, norepinephrine, and epinephrine were also obtained. Finally we wanted to assess the potential for impact on cardiovascular risk. The correlation between serum levels of total cholesterol and heart disease has long been established (Rosenman, Brand, Sholtz, & Friedman, 1976). More recently, however, studies have indicated that the total cholesterol to high density lipoprotein (HDL) ratio may be an even better predicter of cardiovascular risk, especially if this ratio is above 6.0 (Uhl, Troxler, Hickman, & Clark, 1981). Given that stress increases cortisol which controls the production of cholesterol, total cholesterol, HDL cholesterol, and their ratio were also included.

Our goals were to ascertain: 1) Was the stress program associated with changes in the above variables, and 2) Do those at most risk (via cholesterol ratio) differ in their response to stress management.

METHOD

Subjects: Subjects were 26 volunteers from a pool of active duty Air Force and civil service middle managers and secretaries who agreed to participate in a stress management program after an introductory presentation on the influence of stress on cardiovascular disease. There were 15 females and 11 males ranging in age from 21 to 66 years with a mean of 38.3 yrs.

Measures: The following measures were used pre- and post stress management program: STPI with state anxiety (S ANX), state curiosity (S CURI), state anger (S ANG), trait anxiety (T ANX), trait curiosity (T CURI), and trait anger (T ANG) subscales; the Zung Depression Scale; "Your Lifestyle Profile" developed from the pamphlet from the Canadian National Ministry of Health and Welfare; and a checklist of medical complaints covering all systems (MED COMPL). Biochemical data consisted of serum cholesterol (TOT CHOL), high density lipoprotein cholesterol (HDL), and cortisol (SER COR); and urinary cortisol (U COR), norepinephrine (NOREPI), and epinephrine (EPI). The cholesterol ratio (RATIO) was computed by dividing the total serum cholesterol by the high density lipoprotein. Physical measures included systolic (SYS BP) and diastolic (DIA BP) blood pressure and heart rate (PULSE).

Procedure: Subjects had responded to advertisements at Brooks Air Force Base to attend a presentation on executive stress. After the presentation they were offered the opportunity to participate in the study, including the psychological and biochemical assessments.

Before and after the stress management program subjects were provided the packet of the measures listed above. They were also scheduled for an 8:00 am, 12-hour fasting blood sample and pulse check. Blood pressures were taken three times and averaged. Subjects were provided with containers in which to collect an 8-hour urine during duty hours and returned them to the lab after the sample was complete. Late returns received phone prompts.

Cholesterol levels were obtained on an Abbott ABA-100 using the enzymatic method. The cortisol levels were obtained using the radio-immuno assay technique and the catecholamine values employing a radio-enzymatic assay.

Stress Management Program: The stress management program consisted of six, 1-1/2 hour, weekly meetings of the entire group. The program (described in Troxler & Cayton, 1981) was a combination of behavioral, relaxation, communication, and cognitive skills.

Data analysis: The data were analyzed using the UCLA BMDP program to perform paired T-tests. Comparisons were made on the percent change on the psychological and biochemical data by subtracting the post-treatment scores from the pre-treatment scores and dividing by the pre-treatment scores. The T-test was then used to test if the percent change was equal to 0. Change scores on the outcome measures of the High Risk Group (TOT CHOL / HDL > 6.0) and the Low Risk Group (TOT CHOL / HDL < 6.0) were compared.

RESULTS

As indicated in Table--1, completion of the stress management group was associated with statistically significant differences at the p<.01 level or better in urinary norepinephrine, epinephrine, and cortisol and in systolic blood pressures and trait anxiety, all in the therapeutic direction. Differences at the p<.05 level or better were also found in the measures of high density lipoprotein, the total cholesterol/HDL ratio, and diastolic blood pressure, again in the expected direction.

VARIABLE		PRE	POST	STATISTIC	p-value
T ANX	mean	21.80	19.22	t= 3.05	0.0072 **
	s.d.	6.15	5.16	df=17	
MED COMPL	mean	24.08	18.61	t= 1.76	0.096
	s.d.	12.01	9.95	df=17	
PULSE	mean	79.84	75.08	t= 1.90	0.070
	s.d.	11.14	11.14	df=22	
HDL	mean	44.8	48.7	t=-2.28	0.033 *
	s.d.	19.72	15.89	df=21	
RATIO	mean	5.9	4.5	t= 2.42	0.025 *
	s.d.	3.65	1.85	df=21	
U COR	mean	139.7	90.4	t= 3.32	0.003 **
	s.d.	70.76	31.84	df=22	
NOREPI	mean	56.1	41.3	t= 4.58	0.0001 **
	s.d.	20.84	20.70	df=23	
EPI	mean	20.95	15.15	t= 3.92	0.0007 **
	s.d.	8.13	6.06	df=23	
SYS BP	mean	128.37	120.51	t= 3.71	0.001 **
	s.d.	13.95	13.83	df=22	
DIA BP	mean	76.40	73.07	t= 2.28	0.0386 *
	s.d.	11.45	9.40	df=22	

*=p<0.05 **=p<0.01

Table 1-- Measures Whose Pre Post Differences Exceeded or Approached Statistical Significance

Listed in Table --2 are those mean differences in percent change which approach or exceeded statistical significance. The High Risk Group (Total Cholesterol/HDL Ratio > 6) reduced their ratios more than the Low Risk Group (TOT CHOL/HDL RATIO <6), at a probability level of p<.044. In contrast the Low Risk Group reduced their systolic blood pressure significantly more than did the High Risk Group.

VARIABLE		RATIO<6	RATIO>6	STATISTIC	P-VALUE
HDL	mean	-.045	-.903	pool t= 4.16	0.0005 **
	s.d.	0.205	0.785	df= 20	
RATIO	mean	0.048	0.366	pool t=-2.98	0.0075 **
	s.d.	0.188	0.305	df= 6.5	
SYS BP	mean	0.075	0.031	sep t= 1.93	0.068
	s.d.	0.079	0.026	df=19.9	

**=p<0.01

Table 2--Differences in Percent Change Between High Risk-(Chol/HDL Ratio > 6) and Low Risk-(Chol/HDL Ratio <6) Groups Which Approached Statistical Significance

DISCUSSION

The results have obvious limitations in that there was no control group with which to contrast the changes alledged to have occurred as a result of the stress management group. One therefore cannot state that these reductions were caused by the stress program. However, this does not detract from the fact that the values did change, even for those at high risk for heart disease because of their excessive cholesterol to HDL ratios. This is very promising for the fight for reduced incidence of cardiovascular disease and suggests that perhaps these risk factors which have been identified are in fact modifiable on the short run and therefore, perhaps even for the long term. The consistency of the changes across the physical measures of heart rate and blood pressure; the biochemical measures of acute stress, i.e., cortisol, norepinephrine, and epinephrine; the serum marker of high risk, total cholesterol/ HDL ratio; and the psychological measure, trait anxiety is most encouraging. It appears that high risk patients change very much like the lower risk patients except in the areas in which they are at greater risk, in which they change more. This strongly suggests patients with high cholesterol ratios can be helped, at least in the short run and that stress management programs may be a viable approach. The long term effects of course, have yet to be established.

REFERENCES

Allen, J. K. & Blanchard, E. B. Biofeedback-based stress management training with a population of business managers. Biofeedback and Self-Regulation, 1980, 5, 427-438.

Borysenko, J. Z. Behavioral--physiological factors in the development and management of cancer. General Hospital

Psychiatry, 1982, 4, 69-74.
Charlesworth, E. A., Murphy, S., and Beuktler, L. E. Stress management skill for nursing students. Journal of Clinical Psychology, 1981, 37, 284-290.
Dohrenwend, B. S. & Dohrenwend, B. P. (Eds.) Stressful life events: Their nature and effects. New York: John Wiley & Sons, 1974.
Engel, G. L. Psychologic stress, vasodepressor (vasovagal syncope), and sudden death. Annuals of Internal Medicine, 1978, 89, 403-412.
Glass, D. C. Stress, behavior patterns, and coronary disease. American Scientist, 1977, 65, 177-187.
Jenni, M. A. & Wollersheim, J. P. 'Cognitive therapy, stress management training, and the type-a behavior pattern. Cognitive Therapy and Research, 1979, 3, 61-73.
Locke, S. E. Stress, Adaptation, and immunity: Studies in humans. General Hospital Psychiatry, 1982, 4, 49-58.
Lynch, J. J., Paskewitz, D. A., Gimbel, K. S., & Thomas, S. A. Psychological aspects of cardiac arrthymias. American Heart Journal, 1977, 93, 645-657
McQueen, D. V., & Seigrist, J. Social factors in the etiology of chronic disease: A review. Soc Sci Med, 1982, 16, 353-367.
Sklar, L. S. & Anisman, H. Stress and cancer. Psychological Bulletin, 1981, 89, 369-406.
Rosenman, R. H., Brand, R. J., Sholtz, R. I., & Friedman, M. Multivariate prediction of coronary heart disease during 8.5 year follow-up in the western collaborative group study. American Journal of Cardiology, 1976, 37, 903-910.
Spielberger, C. D., Jacobs, G., Crane, R., Russell, S., Westberry, L., Barker, L., Johnson, E., Knight, J., & Marks, E. Preliminary manual for the State/Trait Personality Inventory (STPI). Tampa: University of South Florida, Human Resources Institute, 1979.
Tarnow, J. D. & Silverman, S. W. The psychophysiologic aspects of stress in juvenile diabetes mellitus. International Journal of Psychiatry in Medicine, 1981, 11, 25-44.
Troxler, R. G. & Cayton, T. G. Stress reduction for business and industrial employees. In M. M. Faber & A. M. Reinhardt (Eds.), Promoting health through risk reduction. New York: MacMillan Publishing Co., 1982.
Uhl, G. S., Troxler, R. G., Hickman, F. R., & Clark, D. Relation between high density lipoprotein cholesterol and cononary artery disease in asymptomatic men. American Journal of Cardiology, 1981, 48, 903-910.
Zung, W. A self-rating depression scale. Archives of General Psychiatry, 1965, 12, 63-70.

STRESS MANAGEMENT IN ISRAELI DEFENSE FORCE: IMPLICATIONS FOR U.S. MILITARY HUMAN SERVICE PROFESSIONALS

BY: MAJ FRANK L. GOLDSTEIN Ph.D.
United States Air Force

Few countries in the Western World have devoted as much time and effort to the study of psychological stress and adjustment in war and peace as has the country of Israel. The Israeli Defense Force has monitored the lessons of studies of public responses to terrorism, responses to combat stress and the responses to bereavement and loss.

Specifically, three topics will be addressed that have particular relevance and importance to American military professionals both for our academic knowledge base, and in the adoption and acknowledgement of new responses to an ever increasing stressful world. The three topics in question are: (1) Domestic stress reactions. Particular emphasis being on the effects of having lived through stressful situations, such as terrorist victimization, natural and manmade disasters and the effects of dislocation. (2) Bereavement stress response. Basically, the inclusion of both loss and readjustment in adults to both death and disability. (3) Combat stress reactions. The forward treatment of combat reactions, the use of combat social workers and how to predict who could be candidates for combat type stress reactions.

This presentation does not attempt to be an inclusive study of the lessons that American military professionals can learn from their Israeli counterparts. It does attempt to present a basic outline of the lessons learned in three specific areas over forty years by a typically western nation under constant stress and what American human service practice can draw from that Israeli experience.

DOMESTIC STRESS REACTIONS

Domestic stress reaction may be studied in two parts. First, as it is encountered in a domestic disaster such as a terrorist attack or a volcano. And second, as how individuals react to the stressful event and what provides the most help to those individuals after the event.

A disaster may be considered a state of massive and collective stress which is caused by an external event over which neither the individual nor his society has control. Terrorism may be considered a domestic disaster as it affects the Israeli public. Unlike conventional war, terrorism is a psychological weapon, intended to achieve political ends through intimidation and fear. While actual physical damage does always occur, the main objective of political terrorism is to carry various messages to several audiences at the same time.

With an act of terrorism, there are differential influences on such groups as international public opininon, the inhabitants of the Israeli Administered Territories, the Arab populations and members of the terrorist groups themselves and their likely recruits. The victims of main concern are the target population, the Israeli public. Research indicates that the Israeli public responds to Palestinian terrorism with a high degree of concern (Merari and Keinan, Tel Aviv University 1983). In simple terms, the Israeli reaction indicates or implies success on the part of the terrorist groups since to terrorize the public is a prime terrorist objective. However, how the Israeli public internalizes that stress and their emotional reaction to those acts of terrorism may provide totally different responses than those that would be expected. An Israeli study shows that a hardening of attitudes toward retaliatory measures against the terrorist organizations and concerning the Palestinian issue in general is observed as part of the Israeli problem. This reaction, in turn, has impacted on official Israeli government policy. Thus, the

official Israeli government policy is not always understood when placed against international public opinion.

Additionally, both Israeli and American research has shown that the psychic traumatization to an individual who survives a catastrophic stress situation such as a terrorist attack or a natural disaster (in American terms, the survivors of the Mt. St. Helens catastrophy) is most severe or enhanced when the catastrophy includes the destruction of social and community bonds. The loss of the social structure amplifies the traumatic process and retards recovery from it.

The Israelis have found that the presence of social supportive bonds during severe stress can act as a mitigating influence on the traumatic stress process and protect individual personality functioning against its long term effects. Additionally, social support systems can aid in recovery from the traumatic state. Western academic circles are well aware of the positive use of clinical and non-clinical groups with Holocaust survivors and survivors of natural disasters. However, Western Society is not as well informed on the importance of social support during the event and how crucial it is to post event recovery and rehabilitation. Israeli evidence shows social support protects the victim and assists them in maintaining their social adaptation. This is extremely important in light of recent research that shows that humans who have a depressive attributional style in which they construe failure as stemming from internal, stable and global causes are more susceptible to depression, ill health, and collapse in the face of severe stress. Treatment after the event would be of limited use to these individuals and necessarily of a long term nature.

BEREAVEMENT STRESS RESPONSE

While death and disability have been with us for quite some time, the impact of that loss on segments of our society have been basically overlooked in terms of adjustment to that stress. The Israelis are attempting to predict the factors that enhance and impede readjustment after traumatic loss. They specifically are working with three groups: (1) war widows, (2) bereaved parents and (3) disabled army veterans (Ben-Sira, Hebrew University School of Social Work 1983). Preliminary data has indicated that one's individual resources and primary support group enhance readjustment more than the assistance offered by formal rehabilitative agencies. Also, when an individual has his own resources combined with a primary support group, readjustment is facilitated regardless of the severity of the disability or the lapse of time since the loss. The study is a first step but certainly shows that primary education in emotional support of self, self confidence and personal survival are legitimate roles for helping professionals. The bottom line implication is that those individuals that have a strong personal belief system can survive the traumatic event.

Additionally, there is every indication that traumatic, stress producing events (bereavement in nature) such as loss of life, limb and peace of mind have common dynamics as well as differences. If practitioners can address three specific dimensions: (1) type of loss, (2) stage of coping with the loss and (3) individual strengths and family social system, then those same practitioners have definite guidelines as to which cases will require what type of intervention, the most appropriate time to intervene and finally how long the intervention should last.

The most difficult intervention appears to be that which takes place with the individual who's loss is so global in nature, for example, the entire family is killed or the entire town has had to move and all possessions and sense of community are lost. Many times that individual because of the circumstances, is unable to mourn. That is, the mourning is postponed to a later stage. Israeli studies show that some holocaust survivors have not been able to mourn their losses throughout their lives. At some points natural life events can facilitate this mourning process but for

others only clinical help can act as the catalyst to begin the mourning process. These same studies show that a therapist can be creative and very helpful in working with bereavement victims who are in delayed mourning situations. Examples abound in American 20th century society: Mt. St. Helens, Dioxon contaminated land, flood victims in the south, storm victims in the west and victims of tragic accidents and crime sprees. Many times the survivors have to wait, because bodies are not recovered, towns are missing, support groups are dispersed and professional help is not available till long after the event, a wait which many times is unnecessary and which evidence shows only prolongs the helping intervention process.

A delayed bereavement response presents quite commonly with an individual postponing the mourning during the hectic days, weeks or years following the event only to enter a period of time when the emotional disturbance presents itself during a physical illness, a move, or basically any event that turns out to be another period of stress. The new situation re-evokes hardship and feelings that had not been successfully overcome before. The impression is that we as a society are not allowing those individuals among us with a need to mourn the opportunity. Furthermore, a look at any newspaper will present a full page of national events requiring a period of mourning for stress reduction for those involved. A further look would not indicate bereavement programs being available to those same individuals.

Additionally, while mourning is a normal reaction to loss, its function is to restore an individual's ability to love, which was impaired by the trauma of loss. In deviant mourning, the bereaved is unable to relinquish the loss and the capacity to love another is not restored. In a bereaved family in typical circumstances, a shifting of roles takes place and new relationships compensate for the loss. Family members share the grief and pains and this leads to a strengthening of family ties. In cases of prolonged, deviant mourning a split and deterioration occurs in the bereaved family. Intervention aimed at mending the split is often perceived as siding with some family members against the mourner, and therefore it is resisted. Individual therapy with the mourner may be preferable. The advantage to this short term intervention is that the family conflict is temporarily bypassed and a positive, trusting relationship may facilitate gradual relinquishing of the tie with the deceased (Aleksawdrowiez D.R., Ben Gruion University 1978).

An additional, short term bereavement intervention approach used by the Israelis with the parents of deceased military members was a short term group. A notable addition, in both Israeli individual and short term group treatment, is that a trained volunteer or social worker is in close contact with the family from the time of loss. In the group experience, the volunteer or social worker also accompanied and supported the parents during the first part of their group experience. The groups were open discussion groups.

What the Isareli studies show us is that by and large traditional institutions of society which used to help and support in general no longer fulfill this function in Israeli society (nor in the United States). Therefore, there is a strong need to create a new system to fill the gap. The Israelis feel the medium of short term individual and group work added on to immediate contact after the loss is a positive impact which emphasizes the principals of personal strength and responsibility for one's life. They further feel that mutual help and support are essential pillars of a healthy population. Finally, they see the use of qualified professionals as assuring the attainment of those objectives.

COMBAT STRESS REACTIONS

The Israeli's view "Unit Morale" as the secret weapon of the Israel Defence Forces. While many military commanders see combat stress as a function of poor

morale, the Israeli view is that it is a natural outcome of combat which may be tempered by high unit morale.

Israeli combat stress studies indicate the following: (1) they support American studies of WWII that medical, disciplinary, and psychiatric evacuations analyzed on a day to day basis increased with intensive combat. (2) Support studies that show that psychiatric evacuations were typically preceeded by more intense combat, but disciplinary evacuations were preceeded by more scattered experiences (Noy, Shabtai., Hebrew Univ 1978).

Additionally, Israeli studies showed that typical reactions to combat fatigue such as superstitions and fatalistic thoughts may be interpreted as adjustive reactions and a coping mechanism (Sohlberg, 1975). A further study on the relationship of personality to combat reaction showed little predictive association. However, some predictions did exist. First, the primary factor affecting combat stress appears to be stress outside the combat area (Noy 1978). Another study of civilian background factors such as family, marriage, and education were investigated in a group of 215 soldiers. The investigation attempted to determine what factors were significant to the development of psychiatric disorders during military service. In those individuals not presenting psychiatric disorders prior to military service, stress related to marriage and children were among those variables most significant in psychiatric disorders and not personality.

It is important to note that the most consistant correlates of psychological stress appear to be anxiety. But there are marked individual differences as to what is stressful. It can be said, "that stress and anxiety are both a stimulus situation and a personality trait" (Gregory, W.E., 1978). That is, any one of us will become anxious if the stress is great enough and of the kind to which we are susceptible. Thus individuals may avoid anxiety in stress situations by having stereotyped responses which carry them through their own particular crisis. They would then only develop anxiety if their stereotyped response is interfered with. Logically then, one way of coping with stress is to have a repertoire of behaviors suitable for almost any occasion. Military training and human service training have long advocated this position by the use of practicum instruction. Practicum instruction such as: role playing, war games, flight simulators, parachute jumps, practice crash landings, etc. have a long history of use. The obvious dilemma is that it is difficult to simmulate live combat. However, the Israeli position appears to be that forward treatment of psychiatric casualties and returning them to duty is an extension to the soldiers repertoire of behaviors (Gregory, 1978). Thus the combat stress experience is generally positive as the largest number of psychiatric casualties are not viewed in a negative sense, and are not permanently removed from their comrades. The typical treatment for a soldier who succumbs to combat stress or who develops anxiety in a situation in which he has no stereotyped responses is that he is temporarily removed from duty to a relatively safe area within the combat theater to rest and assimilate his experience. He is assisted by a mental health worker be it a "combat social worker or psychologist" and returned to duty within 24-48 hours whenever possible. For individuals so treated, the recedism rate is close to "Zero" and reactions of both the man and the unit is an increased morale state (Noy, S., I.D.F., 1983).

The Israeli success with the return of stress reaction soldiers to combat creates two significant questions for American human service professionals. The greatest is the duty to increase the repertoire of behaviors to anxiety and stress in our society. The second, and perhaps the most important, point is the therapeutic strategy of quick crisis intervention with limited removal from the offending environment but coupled with support.

SUMMARY

In summary, the Israeli studies provide some very clear guidelines for American human service professionals. The history of the past year provides clear evidence that terrorism as a domestic stress reaction can be expected in the United States despite our relative lack of terroristic acts when compared to a country like Israel. While the actual number of terrorist incidents is generally low, our population is subjected to very intimate media coverage of the incidents. Subsequent security precautions in everyday work and play places make the events totally felt and experienced by our population at large. Additionally, the reactions of Americans to the terrorist may be similar to that of the Israelis. Terrorist attacks may harden the resolve of the civilian population rather than weaken it on any given subject.

For those Americans involved in a domestic stress situation, the need for an immediate support system is well established. The Israeli practice of providing that support system extremely close to the incident is a logical followup to our own knowledge of support, both individual and group following a catastrophic event. With the loss of extended families and the magnitude of some domestic events, the need for helping professionals to support individuals early in the crisis is reinforced by the Israeli experience. The challenge is for us, the helping professionals, to risk meeting the individuals and families at the site and not necessarily waiting in our offices in nearby cities or towns for things to calm down and get back to normal. Just as our medical colleagues respond to medical and domestic disasters so should helping professionals' respond to stress reactions.

In the bereavement stress response the Israeli work with war widows, bereaved parents and disabled veterans is extremely progressive and enlightening. The most significant information they provide is the need for primary education in emotional support of self and the early support of the bereaved. Their specific use of both individual and group therapy methods by qualified professionals at an early point in the bereavement and within a mutual support framework appears very workable.

Combat stress reactions is a historical condition and a natural outcome of combat. The Israeli studies basically reveal much of the information acquired by the United States Armed Forces in WWII. Foremost in their work is the assumption that early return to duty of stressed soldiers is in most cases the preferred approach which provides best results for both the unit and the individual. The Israelis are working on predictive associations to determine which troops are the most likely candidates for combat stress. Their findings, that when all indicators are equal that stress related to marriage and children is the most significant has wide ranging implications for U.S. helping professionals. First, it is apparent that working fathers and mothers do not leave their problems at home when they go to work. For thos individuals in higher stress employments - such as the military, medical services, etc., the cost in training and productivity loss could be significant. Secondly, it further presents a need for ongoing studies of families, family stress and family interrelationships.

The Israeli studies in many ways paralleled some of the work being done in the United States. The work reinforces some work already done and certainly provides areas for future research. The importance of the work presented is that much more can be done for our high stress society, especially in those areas which due to lack of attractiveness have been largely ignored. The current state of the world demands much more attention to these three areas by American human service professional. The work is more important for those of us serving military populations. Not only have we been given insights for working with our troops but very important family issues have been raised which directly impacts on readiness, morale, and ongoing troop support.

Military Leadership Education: What War College Journals Seem To Suggest

David D. Van Fleet
Texas A&M University

Abstract

The new, complex taxonomy of leader behavior suggested by Yukl is used to examine the content of the journals of the War Colleges of the Air Force, Army, and Navy. There are several purposes for such a content analysis. First, if these journals reflect the educational content of the War Colleges, this analysis would reflect that. Second, two time frames are used to examine changes between a period of "war" (1966-1970) and a period of "peace" (1977-1982). Third, these data could form the basis for discussion of numerous points: Is the education relevant for the current time? Is the education relevant for the time when most of "students" would need it? Do the journals reflect the educational content of the War Colleges? And so on.

Introduction

Leadership in general and military leadership in particular have been studied for some time (Bass, 1981; Yukl, 1981). Episodes during the Vietnam conflict (Gabriel and Savage, 1978) and efforts to prepare for futuristic battlefields (Hunt and Blair, forthcoming) have led to recent intensifications of interest in military leadership. Hopefully, however, this new round of research will not repeat the problems of the past, but rather will build upon the successes of previous work to provide a clearer understanding of military leadership for the future.

Specifically, the level of analysis in emerging research must be intermediate. The highly generalizable, but not very useful, task- and maintenance-oriented dichotomy is at too high a level of abstraction to be of much practical significance and more research based on such a dichotomy will not extend our knowledge very much. If one or more subcomponents of one or both of those broad categories is vital while another is not, research using such dichotomies will likely fail to uncover that vital component. On the other hand, the use of highly specific job and position analyses based on detailed behavior or competencies may also not be so useful. Those analyses are conducted for particular positions at particular organizational levels in particular organizations at particular points in time. As such, they may be

quite useful for specific, short-term job training, but not be so useful for more general leadership education. They cannot be generalized very well. Intermediate level analysis is needed.

Intermediate level analysis would be much more detailed, involve many more dimensions than higher level analysis, but those dimensions would be generalizable across a moderately wide range of circumstances. A new taxonomy of leader behavior has been developed for just such analysis (Yukl, 1982). For more than six years, careful, systematic research has been conducted and has led to the development of this new taxonomy. The taxonomy appears broad enough to capture most relevant leader behaviors and yet is quite useful in specific applications as well.

In other research, this new taxonomy has been demonstrated to be both applicable to and useful in research on military leader effectiveness (Yukl and Van Fleet, 1982). That research involved using the taxonomy with both critical incident and questionnaire-correlational methodologies. Another obvious application of the new taxonomy would be in content analysis research. A variety of specific steps might be employed in content analysis, but fundamentally the content of something is examined to see to what extent that content can be represented or described by the categories in the taxonomy. That "something" could be interview protocals, diaries, reports of observers, or other written material such as books and journals.

Purpose

The purposes of this paper are (1) to demonstrate that the new taxonomy can be used in content analysis research and (2) to determine what can be learned from such an analysis on material written about military leadership and used in the education of military officers.

Method

On the assumption that the journals of the several War Colleges reflect, at least in part, the content of the education provided by those War Colleges, those journals were selected for study. Thus, three journals were content analyzed--the Air University Review (AUR), the Military Review (MR), and the Naval War College Review (NWCR). To determine if that content has changed over time, two particular five-year time periods were selected--one of relatively low conflict, 1977-1982, which will be termed a "peace" period and one of higher conflict, 1966-1970, which will be termed a "war" period. Five-year time periods were used to "smooth" short-term fluctuations in content.

For each of these five-year time periods, a graduate student familiar with the definitions used in the taxonomy performed the analysis on all articles with management or leadership (or related terms) in the title or abstract. These were done for each such article in each issue of each journal for the two five-year time periods. As this procedure is extremely time consuming, multiple analysts were not used. However, the accuracy of the procedure was checked by having the judgment of one of these analysts verified by having a group of six other graduate students content analyze various issues of the 1982 Military Review. The +0.81 correlation obtained suggested that researchers familiar with the taxonomy can perform consistent analyses of journal content using that taxonomy. The specific analysis involved the

student reading each article and making a tic mark next to a category from the taxonomy every time he or she read something which dealt with the content of that category.

Results

The results of this content analysis may be viewed in either of two ways--across time or across the journals. The former approach would suggest either "war" or "peace" differences or at least developmental differences over time. The latter approach would suggest differences across service branches or perhaps across editorial review boards. Each of these is examined below based on the data shown in Figure 1.

Insert Figure 1 About Here

Before looking to those differences, however, a comment is in order about one striking piece of information in this figure--the average number of tallies per issue. In each time period the number of tallies for the Military Review is greater than for the other two journals and for the peacetime period it is phenomenally greater! No simple reason is evident for the wartime difference, but for the peacetime difference one does exist. The Military Review ran a special issue on leadership during this time period. That issue itself was a source of many of the added tallies, and it also generated other articles, letters, etc. which increased the tally count for that journal. This wide discrepancy in tallies is one reason why the data in the figure are in percentage terms. But just what do those data suggest?

For the Military Review, only two of the 23 categories are significantly different across the two time periods. Consideration and Facilitating Cooperation and Teamwork are both greater in times of relative peace than in time of war. For the Air University Review, seven categories are significantly different. Clarifying Work Roles is less frequent in peace while the other six are all greater. Those six are: Goal Setting, Emphasizing Performance, Training-Coaching, Problem Solving, Disseminating Information, and Criticizing. For the Naval War College Review, 12 of the 23 categories are significantly different. Five were greater in peacetime (Inspiring Subordinates, Problem Solving, Showing Consideration, Monitoring Operations, and Delegating), while seven were greater in wartime (Managing Conflict, Administering Discipline, Clarifying Work Roles, Planning, Goal Setting, Training-Coaching, and Disseminating Information).

A closer look at these indicates that in no case was a single category significantly different for each of the three journals. In only six cases was a significant difference obtained for two of the three journals for a single category, and the direction of difference was alike in only half of those! Do these differences reflect genuine differences among the branches of the service in terms of their missions and the leader behaviors necessary to accomplish those different missions? Or are they reflecting differences in philosophy which may or may not actually be related to leader effectiveness? Or are they merely random or something else? Further, are these differences what one would predict based on leadership theory? Certainly the Showing

Consideration and Clarifying Work Roles differences are in the direction which one might expect based on theory, but Problem Solving seems reversed. Since other differences vary across the journals, there is little to be gained from further speculation at this time. More research will be necessary before we can reasonably begin to fully discuss this differences, but this is a crucial area in need of further and careful research.

A look across journals reflects a similar pattern. The Military Review is different from the others in only two instances--the peacetime frequencies are different for Problem Solving (lower) and Facilitating Cooperation and Teamwork (higher). Air University Review has four instances of differences, again all peacetime and again some greater (Goal Setting, Disseminating Information, and Criticizing) and one less (Showing Consideration). Finally, the Naval War College Review has eight different instances and some are for wartime while others are for peacetime. All of the wartime differences (Managing Conflict, Administering Discipline, Clarifying Work Roles, and Disseminating Information) and three of the four peacetime (Monitoring the Environment, Monitoring Operations, and Delegating) are greater. Only the one remaining peacetime difference (Training-Coaching) is less than that of the other journals. Again one wonders why these differences exist, and again there is no ready answer.

Conclusion

The primary purpose of this paper was to demonstrate that Yukl's new taxonomy of leader behavior could be effectively utilized in content analysis research, specifically for content analyzing journal articles. That purpose has been clearly accomplished. A second purpose, however, was to determine what can be learned from such an analysis. We have learned that there are war versus peace differences and differences among the services. We have not learned, however, anything about why such differences exist.

References

Bass, B. M. Stogdill's handbook of leadership. New York: The Free Press, 1981.

Gabriel, R. A., & Savage, P. L. Crisis in command. New York: Hill & Wang, 1978.

Hunt, J. G., & Blair, J. R. Leadership in the Air-Land Battle 2000. Forthcoming.

Yukl, G. A. Leadership in organizations. Englewood-Cliffs, New Jersey: Prentice-Hall, 1981.

Yukl, G. A. Innovations in research on leadership behavior. Paper presented at the Eastern Academy of Management, Baltimore, Maryland, May, 1982.

Yukl, G. A., & Van Fleet, D. D. Cross-sectional, multi-method research on military leader effectiveness. Organizational Behavior and Human Performance. 1982, 30, 87-108.

Figure 1

Content Analysis of Selected Military Journals
(Percentage Frequencies Rounded to Nearest Hundreth)

Behavior Categories	War Years (1966-1970) Military Review	War Years (1966-1970) Air University Review	War Years (1966-1970) Naval War College Review	Peace Years (1977-1982) Military Review	Peace Years (1977-1982) Air University Review	Peace Years (1977-1982) Naval War College Review
Managing Conflict	.02	.03	.08 a,b,c	.03	.02	.00 b
Administering Discipline	.02	.02	.09 a,b,c	.03	.00	.00 b
Planning	.07 a	.05	.06 b	.04	.00	.00 b
Goal Setting	.04	.05 b	.08 a,b	.05	.15 a,b,c	.02 b
Clarifying Work Roles	.07 a	.08 a,b	.12 a,b,c	.03	.02 b	.00 b
Inspiring Subordinates	.06 a	.07 a	.03 b	.09 a	.05	.10 a,b
Emphasizing Performance	.08 a	.08 a,b	.06	.09 a	.03 b	.07 a
Training-Coaching	.06 a	.03 b	.10 a,b	.07 a	.08 a,b	.03 b,c
Showing Consideration	.04 b	.05	.05 b	.11 a,b	.07 a, c	.17 a,b
Problem Solving	.03	.03 b	.01 b	.02 c	.09 a,b	.10 a,b
Disseminating Information	.04	.03 b	.11 a,b,c	.04	.10 a,b,c	.05 b
Delegating	.03	.03	.03 b	.02	.00	.10 a,b,c
Encouraging Decision Participation	.03	.03	.03	.05	.07 a	.05
Criticizing	.03	.01 b	.03	.05	.19 a,b,c	.00
Facilitating Cooperation and Teamwork	.04 b	.04	.00	.08 a,b,c	.01	.00
Monitoring Operations	.02	.03	.00 b	.04	.01	.17 a,b,c
Providing Praise and Recognition	.04	.06 a	.05	.02	.01	.03
Career Counseling	.02	.05	.01	.03	.01	.00
Structuring Reward Contingencies	.04	.06 a	.01	.04	.00	.00
Facilitating the Work	.04	.03	.03	.03	.03	.03
Innovating	.04	.03	.00	.02	.03	.03
Representing the Unit	.02	.03	.00	.03	.00	.00
Average Number of Tallies per Issue	3.76	1.98	2.16	14.98	1.43	2.00

a = highest frequencies

b = war/peace significantly different ($p < .05$)

c = proportion significantly different from each of the other two journals ($p < .05$)

AD-P003 303

The Role of Relevant Experience and Intellectual Ability in Determining the Performance of Military Leaders: A Contingency Model Explanation

Patrick J. Bettin
United States Military Academy

Abstract

A field study involving 79 army combat officers in middle echelon leadership positions was conducted to evaluate the role of relevant experience and intellectual ability in predicting leadership performance. Biographical and organizational data were the primary measures used to ascertain the relevance of leader experience. This represents a departure from previously used methodology which considered only the leaders' organizational tenure in determining experience levels. Results shed light on the components of Fiedler's Contingency Model of Leadership Effectiveness supporting the hypothesis that task- and relationship-motivated leaders make effective use of their experience only in situations which match their leadership personality. The study also suggests a plausible relationship between the leader's cognitive resources and leader behaviors.

There is a widespread belief that experienced leaders bring job relevant skills, knowledge, and judgment to their positions. The term, experience, is defined in Webster's dictionary as "...knowledge, skill, or practice derived from participation in, or direct observation of an activity..." (Gove, 1971). This definition presupposes that as an individual gains experience, he or she also gains knowldege or skills which are then transferable to accomplish subsequent tasks. Experience, in other words, is generally viewed as a valued resource to be brought to the new position or acquired during the leader's tenure. It is one of the most widely used predictors in judging an individual's suitability for a job or for a promotion.

Although experience is generally thought to enhance performance, the empirical evidence does not support this view. In a report of findings from three experimental tests and seven field studies, Fiedler (1970) failed to find a signficiant relationship between years of organizational service and leadershp performance. These results (median correlation = -.12) indicated that length of service in an organization did not contribute positively to group performance.

In an effort to explain these counter-intuitive findings, Fiedler and his associates investigated the moderating effects of the situation on the use of leadership experience. Their research concentrated on the effects of interpersonal stress between the leader and his or her superior. This study extends that research and investigates the role of situational favorableness as it affects task- and relationship-motivated leaders differentially.

Method

Subjects. This study investigated the effects of experience on the performance of middle-echelon leaders of an Army infantry division. It involves the company commanders and staff officers of ten combat battalions. These subjects were drawn from a population of army officers occupying positions which normally call for the rank of captain. Restricting the sample to this grade makes it much less likely that experience is confounded with previous performance. A total of 84 officers participated in the study: 44 company commanders and 40 battalion staff officers. With the exception of seven lieutenants serving as staff officers, all subjects held the rank of captain.

Tests and Measures

Experience. Researchers have most commonly defined experience as time in a particular organization, time in a specific position, or time in a general occupational field. As a main effect, these measures have not yielded a consistent relationshp between years of "experience" in an organization and managerial or leadership effectiveness. These surprising, non-obvious results suggest that time may not be an adequate measure of a leader's experience.

In essence, a person's experience level encompasses what he or she has done, where the person has done it, and for how long. Although experience is routinely discussed in terms of time spent doing something, it is important to realize that experience is a psychological, not a physical, dimension. Even though time may be necessary to gain experience, it may not be sufficient to capture the sum total of knowledge and skills a person acquires. Time, therefore, should not be considered the sole measure of experience.

Each subject in this study completed a detailed biography of his work and training experiences since entering the army. Independent, expert judges, blind to the identity of the subjects, assessed each officer's experience level. The interrater correlation, a measure of the reliabiity of the method, was.81 ($p < .001$).

Intelligence. A modified version of Horn's Crystallized (Gc) and Fluid (Gf) Intelligence Sampler (Horn, 1968) was administered. Crystallized intelligence has been related to scores on complex tasks and common cultural activities. It is correlated with formal education and seems to measure what a person has learned over time. Fluid intelligence, or on the other hand, correlates with speed of learning in novel situations. It appears to measure person's ability to solve problems which are unique or have not been previously learned. The two intelligence measures, Gc and Gf, are relatively independent, with inter-correlations ranging from .17 to .53 and test-retest reliabilities between .75 and .87 (Horn, 1968).

Leader Performance. The performance of each subject was evaluated by the superiors in his immediate chain of command. The raters were asked to complete a 49 item, 8-point Likert type scale on each subject. The scale was designed to measure effectiveness in such areas as task performance, communication, subordinate development, decision making, and interpersonal relations. Scores were standardized for each rater permitting comparisons among officers from different organizations. The substantial agreement between raters ($r = .62$, $p < .001$) permitted the ratings to be summed. The overall, standardized performance score had a high degree of reliability (Cronbach's alpha - .94).

Leadership Style. The leadership motivation of the subjects in this sample was measured by the Least Preferred Co-worker (LPC) scale (Fiedler, Chemers, and Mahar, 1976). The LPC score is interpreted as a measure of the leader's motivational hierarchy. Leaders with relatively low LPC scores are considered to be task-motivated while those with high LPC scores are considered to be relationship-motivated. The latter primarily desire a cohesive, pleasant group. Therefore, interpersonal relations receive the leader's attention when the situation is tense or his relations with the group members seem to be tenuous. However, when the goals of being related are satisfied, the high LPC leader shows concern for the task relevant aspects of the group's activities. Conversely, the major objective of the task-motivated leader is to accomplish the task and thereby earn self-esteem by doing a good job. When this need is being satisifed, he seeks friendly, good interpersonal relations with his co-workers (Fiedler, 1972).

In this sample, LPC scores ranged from 18 to 108 with a mean of 61.23 and a standard deviation of 17.15. The sample was divided into two groups; those scoring 63 and below were considered to be primarily task-motivated, those scoring 74 or higher were classified as relationship-motivated. This split resulted in a loss of 5 subjects with LPC scores between these standard cutoff points.

Situational Control. Fiedler's Contingency Model postulates that the performance of leaders is dependent upon two interacting factors, the leader's style and the situational control. Situational control is the degree to which the situation provides the leader with influence, control, and power. This research investigated the impact of role ambiguity, uncertainty, and stress on the leader's perception of situational control.

Stress. The degree of leader-perceived job stress was assessed by a 23 item scale designed to distinguish stress from five theoretical job dimensions: role conflict, role ambiguity, organizational decision-making, interpersonal competition, and unethical competition. Contrary to the a priori attempt to develop five distinct stressors, this scale had a reliability coefficient of .92 ($p < .001$), indicating it is unidimensional and quite reliable.

Results

The preception of task structure is positively related to the leader's experience level and negatively related to the leader's intelligence ($r=.26$, $p < .05$ and $r = -.33$, $p < .01$ respectively). More experienced leaders view their responsibilities as less complex, more structured, and are able to determine when they have been properly accomplished. More intelligent leaders, on the other hand, seem to interpret their duties as being more complicated, less precise, and having more than one correct solution. By seeing more facets of the problem and more alternative solutions, the intelligent leader is confronted with situations which they believe require more information and resources.

More intelligent leaders perceive a more stressful relationship with the superior ($r =.29$, $p<.01$) and more stress resulting from their job ($r =.26$, $p <.05$). This may be interpreted by the notion of cognitive complexity. More intelligent leaders view their environment and interpersonal relationships as more complex. For these leaders, there are far fewer "black and white" situations and many more circumstances in which the solution depends upon a number of alternatives.

The interrelation between stress and situational control variables is quite interesting and provides support for the concept posited by Fiedler (1982) that stress reduces a leader's control over his situation. Leaders who report good relations with their subordinates report less stress resulting from thier job than do leaders with poor leader-member relations ($r = -.33$, $p < .01$). This makes perfectly good sense if one takes the position that discharging one's duties through the actions of subordinates is a primary aspect of leadership. If you trust your subordinates it is less likely that you will view your job as stressful. In addition, there is a significant negative relationship between task structure and job stress ($r = .28$, $p < .05$) and between task structure and boss stress ($r = -.26$, $p < .05$) which permit alternative explanations. Leaders who perceive their jobs as complex also view their work environment as more stressful. Thus, complex jobs may contribute to the perception of stress or stress on the job makes one's leadership situation seem less certain and therefore more complex. Whatever the explanation, in this sample, at least, situational control was related to the perceived stressfulness of the environment. It is very important to note that job stress is not related to rated leadership performance. Equally important, the stress level was not a result of the performance evaluation.

Leader-member relations are unrelated to the leader's experience level. Having experience is not a prerequisite for having good relations with one's subordinates. Although one would expect more experience leaders to have more power, authority, and influence, that does not appear to be the case in this sample. Power in the military is significantly related to the leader's official position. Commanders reported significantly more position power than do the staff officers ($F = 19.79$, p. $< .001$).

Since we are relatively certain that leader behaviors are a result of the interaction of the leader's personality and the situation, it seems that the next logical step is to investigate the impact of the situation on the use of experience and intelligence by task- and relationship-motivated leaders. This will provide clues as to why the more experienced leaders excel in specific situations but not in others. It was hypothesized that relevant experience will correlate with leadership performance for task-motivated leaders in high control situations and for relationship-motivated leaders in moderate control situations. In other words, in situations which "match" the leader's style, the leader is able to take advantage his previously learned skills and abilities. The date from this study provide ample support to the concept that leader's use their experience in situations which match their style.

According to the Contingency Model, task-motivated leaders perform best in situations of high control. As predicted above, it is in this situation that the task-motivated leaders' experience correlates significantly with rated leadership performance ($r = .52$, $p < .01$). Conversely, the high LPC leaders performed poorly in this situation. Note that it is in this situation that their experience is not related to performance ($r = -.07$, n.s.). In moderate situations, the relationship-motivated leaders performed best. And, as hypothesized, in this moderate control situation the high LPC leaders' experience was positively related to performance ($r = .45$, $p < .05$) while the low LPC, task-motivated, leaders' experience was unrelated ($r = .10$, n.s.).

Conclusion

As stated earlier, the major contribution of this study is the finding that task- and relationship-motivated leaders make effective use of their experience in situations which match their leadership personality and misuse their intelligence in situations which do not.

The results of this study support the Contingency Model and provide new insight into the reasons why certain leaders excel in some situations and perform poorly in others. Research has shown that, although low LPC leaders are primarily motivated by task accomplishment, when they are in high control situations they are able to focus on their secondary motivation which is good interpersonal relations with the group members. As a result, in these situations task-motivated leaders behave in a considerate, supportive manner. In moderate situations, the high LPC leaders emphasize their primary motivation, need for relationships. Therefore, they, too, display open, considerate, participative behaviors in situations which match their leadership style.

The major theoretical implication of this information is that we may be able to make a link between cognitive resources, experience and itelligence, and leader behavior. In circumstances in which the leader behaves in a considerate manner leaders tend to use their experience. It seems that experience may be related to interpersonal activities. Experience may tell us how to treat others and act interpersonally. Intelligence, on the other hand, is significantly related in situations in which the leader tends to focus on the task relevant aspects of the jcb and ignores the interpersonal aspects. For the low LPC leader, this occurs in the moderate situation. For the relationship-motivated leader, it occurs in the high control situations. In conclusion, it appears that experience is related to interpersonal activities and intelligence is related to task activities.

References

Fiedler, F.E. Leadership experience and leader performance — another hpothesis shot to hell. Organizational Behavior and Human Performance, 1970, 5, 1-14.

Fiedler, F.E. Predicting the effects of leadership training and experience from the contingency model. Journal of Applied Psychology, 1972, 56, 114-119.

Fiedler, F.E. The effects of leadership training and experience: a contingency model interpretation. Administrative Science Quarterly, 1972, 17, 453-470.

Fiedler, F.E., Chemers, M.M., and Mahar, L. Improving leadership effectiveness; the leader-match concept. New York: John Wiley and Sons, 1976.

Fiedler, F.E., Potter, E.H., Zais, M.M., and Knowlton W.A. Organizational stress and the use and misuse of managerial intelligence and experience. Journal of Applied Psychology, 1979, 64, 635-647.

Gove, P.E. (ed) Webster's Seventh New Collegiate Dictionary. Springfield, Mass: G. & C. Merriam Company, 1971.

Horn, J.L. Organization of abilities and the development of intelligence. Psychological Review, 1968, 75, 242-259.

AD-P003 304

There is No Such Thing as Leadership!

Charles Jacob Yoos II
United States Air Force Academy

Abstract

In this paper, I present the radical but not nihilistic perspective that leadership is an invalid construct. More particularly, I argue that leadership science is a contradiction in terms, and that leadership is at best an *ex post facto* notion, a tenet of faith akin to that held by the ancient alchemists. The purpose of the arguments is not to debunk leadership per se, but to show that clinging tenaciously to the euphemism leadership is counterproductive to the vital pragmatic task of understanding and regulating human behavior.

Introduction and Caveats

I know, I know . . . how dare I commit such blasphemy? Some 500 years after Martin Luther, I tack this thesis to the door of the mighty fortress called military leadership. Like Luther, *Hier steh' ich, ich kann nicht anders!*[1]

Before I start, you ought to be aware of some things:

-- Despite my affiliation with the academic discipline of management, this paper is not about leadership versus management. That is a specious debate anyway, but it is not the subject here.

-- This is also not about leadership as art versus science. My arguments cut deeper than that distinction.

-- I concede the use of the term "leadership" as a gross *rubric*, in general if ambiguous lay usage. I object to the use of the term *as if it captured knowledge*.

-- I observe the empirical legitimacy of much of the voluminous behavioral research which appears under the title of leadership, and I acknowledge the powerful scholarly pragmatic which keeps it there. I deny, however, that this situation gives meaning to the title.

-- I sometimes devil's advocate, but not here. This paper does not reflect the official views of the Air Force Academy or the Department of the Air Force, but it does reflect mine.

---

[1] Here I stand, I cannot do otherwise.

Cogito, Ergo Sum

How can I dismiss the notion of leadership in full view of an awesome literature, including such scholarly tomes as Ralph Stogdill's seminal compendium, Handbook of Leadership: A Survey of Theory and Research (1974), James MacGregor Burns' Pulitzer Prize winning Leadership (1978), and in the military context, Bob Taylor and Bill Rosenbach's forthcoming Military Leadership: In Pursuit of Excellence, as well as countless journal articles and other published treatments?

I do so by proceeding from the Cartesian apothegm which is the heading of this section. Conceptualization must precede existence; an idea must be had before it is labeled. It is ludicrous to do otherwise in the pursuit of knowledge. As an aspect of science, ideas are generated purposefully to be vehicles for understanding; in that these ideas are literally "constructed," they are called *constructs*.

In addition to the necessity of there first being an idea which gives existence to a construct, criteria are also needed for accepting or rejecting candidate constructs, for deciding which constructs we ought to try to validate, and which we ought to discard. Three first-order criteria that are germane here are that a construct should be *explanatory* (it should answer, or at least give insight into, the questions that we ask), *coherent* (not mysterious), and *unique* (not redundant with other constructs). A second-order criteria, by which I mean one that would eventually have to be satisfied, but not until after the first-order criteria are met, is the canon of *intersubjectivity*--the meaning of a construct must be agreed to by the community of inquirers. Humpty Dumpty need not apply!

Within this framework, consider leadership as a construct. I shall pose two versions of my argument: the Devoid Condition, and Occam's Complaint.

The Devoid Condition

I find that leadership is an icon so venerated by time and anecdote that it has assumed the status of a construct without ever satisfying the requisite criteria. Gertrude Stein's dismissal of the city of Oakland is apropos of leadership: there is no "there there." The bulk of contemporary discourse on leadership commonly takes as *a priori* the existence of such a construct, the meaning of which must be searched for until found. This is exactly *sum, ergo cogito,* and it is absurd.

Candidate constructs abound, but they violate even the sparse criterion set that I specified above. I cannot possibly treat these exhaustively in a paper of this scope, but I will offer two generalizations. First, many of them fail because they are not in any way unique--they offer no ideas not already captured in more basic constructs. Second, the preponderance of the empirical research defines leadership as (certain) behaviors of the incumbent of a position designated a leadership position, or defines a leader as one who exhibits leadership behaviors with significant frequency. Such research is typically a methodologically sound crack at the vexing nut of understanding human behavior in its myriad contexts, and I applaud it in its own right, but

no amount of it justifies the claim that there is in fact something called leadership; from a conceptual standpoint, these definitions are trivial if not tautologous. What is more, the variety of definitions which preface this body of research grossly violates my second-order condition--there simply is no substantial agreement about the meaning of the term. Some authors even plead that leadership is such a multi-faceted notion that it is impossible to articulate, bringing us full gamut from the merely ridiculous to the "ineffably sublime!"

Let me illustrate further by analogy. Suppose that you and I are on a balcony, several stories above the ground, and I tell you that there is a physical scientific construct, called "derf,"[2] that enables you to leap over the railing, glide gently to earth, and land safely. I point out that most of us have seen this happen, perhaps in the movies, or at least have heard about it. Naturally, you inquire as to the particulars of derf. I respond soothingly that many derf engineers have defined derf in different ways, and others have admitted that derf is too complicated to explain, but not to worry, derf surely exists. Then, I invite you to vault the balcony railing, chanting "in derf we trust."

The point is not that you would refuse to go over the side (you would refuse, wouldn't you?). Rather, it is that you would not even admit the construct called derf. You might respond, in exasperation, "That's nonsense! Go get some knowledge, about levitation or molecular transportation ('Beam me up, Scottie!') or something, and then we'll label it, explore it, and maybe even test it." I agree, and suggest that the same admonition applies to the social nebula called leadership.

Occam's Complaint

Let's contemplate the same problem from another direction. Recall that Occam's Razor calls for parsimony--explanation is sought with as few concepts as possible, and redundancy is abhorrent. With this in mind, the question is: what knowledge would be foregone, what insight lost, if the term "leadership" were expunged? Go beyond semantics--what would we not know then that we know now? For example, armed with the observation that people seem to effect other people (which we might label *social influence*), the idea that people's capacity to influence, or latent influence, can be thought about and described (which we might label *power*), the ancillary idea that there is a peculiar sort of power which is organizationally legitimated (which we might label *authority*), some codification of right and wrong (which we label *morality*), and a model which processes these ideas, what new deposit does leadership make in our fund of knowledge? Otherwise, leadership is bankrupt.

Phenomenon versus Outcome

Well, if leadership is not a construct, is it still a *phenomenon*, worthy of theorization? A phenomenon is something that happens. Does leadership happen?

---

[2] I had a classmate named Fred who tended to do things backwards, so he was dubbed "Derf."

I am not as zealous on this point, but I must observe that in so far as *precision* and *consensus* are vital aspects of identifying phenomena, then leadership does not qualify here either. At the extremes, we may get broad agreement that Churchill (or Roosevelt or Stalin) exhibited leadership, and that a gun to the head in a dark alley is not an instance of leadership, but there is much ambiguity in the broad expanse in-between, and, in any case, nobody is able to articulate why it was or was not leadership. The French idiom poses it best: *Qu'est-ce que c'est que ça*--what is it that it is? Until the "that" can be specified and agreed to, leadership cannot be considered a phenomenon.

The problem arises because *outcome* is mistaken for phenomenon. Things do happen, after all: military battles are won, businesses realize profit and growth, etc., and people called "leaders" seem to be pivotal in all this. These are outcomes, however, and the error comes in succumbing to the tendency to account for some or all of the outcome by ascribing it to something called leadership, when you cannot say what that is. It would be equally valid to ascribe it to a roll of the cosmic die, or perhaps Divine Intervention . . .

On God & Love

At this point, the tenacious proponent of leadership might say, "Aha, what about that? Don't you believe in anything that you can't feel, see, or otherwise 'know' in a strict sense. What about *God*? What about *love*? Don't we hold these ideas without verification? Why is it OK for them, but not for leadership?" At the risk of not doing justice to these topics, I will treat them briefly in turn.

I believe in God, and it is precisely the point that no matter how great my fervor, that belief is still a *tenet of faith* in my scheme of knowing. I have made a conscious decision--I "choose" to believe in God. I am compelled to do so, for complex, personal reasons, and I accept that God "passeth all understanding." This is my religion. On the other hand, no physical or social construct which I may devise or embrace to help understand this world is part of my religion. If one confirms the analogy between the idea of God and the idea of leadership, then it seems to follow that leadership, like God, is a tenet of faith in your belief system. That can't be refuted, but it ought to be recognized and admitted, and it is also prudent to recall the fate of alchemists, who held as a tenet of faith that base metals could be changed to gold.

As for love, I find that it is a reasonably precise phenomenon about which broad consensus exists as far back as the ancient Greeks, who wrote about such types as *eros*, *filios*, and *agape*. Love may even have utility as a construct--there are certainly approaches that are explanatory, coherent, and unique--consider Freud (1933), Fromm (1956), and May (1970). Each offers a conception that carves out fresh territory of understanding. I reach no conclusion about them, but observe that on this basis, love and leadership are incomparable.

Reconciliation

Is the Emperor called leadership naked? I have argued that it is, but I also think that all such arguments subordinate to the following utterly pragmatic consideration: given this amorphous label called leadership, are we on balance better off with it in our knowledge domain than without it? In this sense, the assertion that "there is no such thing as leadership" is a *non sequitur*--of course there isn't! All constructs are figments, albeit purposeful ones, of our collective imagination. The important question is not whether there is leadership, but whether it is advantageous to (continue to) pretend that there is.

I conclude that we are not making headway on the leadership tack. I read too much literature that promises analysis based on leadership, but produces only artifactual descriptions or empirical generalizations. I hear too many senior executives who speak *ex cathedra* about leadership, but are really prisoners of their own experience. Most grievously by far, I encounter too many students who have been titillated with visions of grandeur about a grail called leadership, but who are confused and frustrated because its mystery has never been revealed unto them, and who resist learning the rigorous fundaments of social psychology because "leadership is where it's at!"

I also suffer no delusions--I cannot prevail on a diet of worms. In addition to stating my case, I can only offer advice. With all deference to those who conduct valid inquiry into human behavior which may reside under the rubric of leadership but does not depend on it, be vigilant that the convenient label not be construed as incorporating meaning. To those who presume leadership, . . . , take stock.

References

Burns, James MacGregor. *Leadership*. New York: Harper & Row, 1978.

Freud, Sigmund. *New Introductory Lectures on Psychoanalysis*. New York: W. W. Norton, 1933.

Fromm, Erich. *The Art of Loving*. New York: Harper & Row, 1956.

May, Rollo. *Love and Will*. New York: W. W. Norton, 1970.

Stogdill, Ralph. *Handbook of Leadership: A Survey of Theory and Research*. New York: The Free Press, 1974.

Taylor, Robert L., and William E. Rosenbach (eds.). *Military Leadership: In Pursuit of Excellence*. Boulder: Westview Press, forthcoming.

Productivity, An Organizational Rather Than An Individual Variable[1]

William D. Kane, Jr.
Western Carolina University

Abstract

This research investigated the impact of environmental distractions on maintenance productivity. The investigator spent 107 hours observing subjects and concluded that over fifty percent of their work shift was spent coping with distractions that hindered task accomplishment. The data suggest that the relationship between productivity and organizational variables is critical.

Introduction

The objective of this research was to conceptually determine whether or not an individual technician's response to contextual variables is a significant component, along with task accomplishment, of performance and/or productivity. Previous investigators purposefully strove to control contextual variables using a laboratory or semi-controlled environment focused completely on the technological components of the task. However, when technicians function in a shop, flight line, or silo environment, contextual variables may play a substantial role in total task completion. Noise, weather, and lighting conditions are recognized as impacting on performance, but no research exists on how technicians cope with or adapt to environmental variables. What should be included as contextual variables of the maintenance environment is a research question, how technicians cope with or respond to them are pertinent to performance, yet no research currently exists which investigates the impact of these contextual variables on performance. It may be that a technician's coping with environmental distractions is as important to productivity as technical task accomplishment.

To investigate the hypothetical relationships, the author spent 107 hours on three different Air Force bases directly observing, in their work environments, eight crew chiefs and nine specialists. The suspected relationships are there. They are more important to productivity than originally thought, but the research problem is far more complex than anticipated. The aircraft maintenance environment is not a monolithic one but is instead a multiplicity of task environments. As the missions of MAC, SAC, and TAC differ, so do their maintenance environments. The environment of a specialist working in a quiet shop is markedly different from the environment of the same specialist working on a noisy, busy flight line. What else has emerged from this research is reinforcement for the idea that maintenance technicians are exposed to an almost infinite queue of environmental demands from which they must select, using some kind of a selection system, and available time (or other resources) constrains the number of events they can attend to. Therefore, they select some, neglect others and the select/neglect activity is not necessarily according to conventional wisdom (i.e., maintenance

---

[1]This research was funded by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under minigrant number AFOSR-80-0146.

policy). While this line of research is far from complete it does indicate some interesting ways to seek increased productivity.

This piece of research was proposed as a conceptual study and as such cannot be generalized to the maintenance population in its current form. Site selection was not random, subject selection was not random, the number of subjects is quite small, sophisticated statistical techniques cannot be used, and there was only one observer.

Observation Methodology

The central purpose of this research was to observe Air Force Maintenance personnel as they went about their duties in as naturalistic a manner as possible. Realizing that the observer becomes part of, or perhaps intrudes into, the maintenance environment it was still believed that after some minimal period the observer would blend into the environment if he were unobtrusive enough. After carefully explaining that the observer was not a quality control inspector, not a manpower specialist and that he would not report performance information to anyone, he was well received. The observer then informed the subject that he just wanted to observe how things got done and that he would accompany the subject but the subject was to go about his or her duties as if the observer was not there.

Results

Base one. The methodology was in trouble from the first hour of observation. The methodology was far too simplistic to observe, record, and interpret the richness and variety of activities and interactions in the environment. The results from base one are divided into two categories and they are treated separately.

A. Results Pertinent to Methodology--As the original methodology was applied to the research task the following problems were encountered and noted.

1. It is oftentimes difficult to determine precisely what the assigned task is and what the parameters are.
2. It is difficult to determine from observation the impact of some environmental interactions.
3. The observation time of two hours was too brief to observe a representative sample of the subject's behavior.
4. The comparison of coping behavior between a high performer and a low performer could not be made for several reasons:
    a. In most instances there were not enough people available for duty to have the luxury of picking between specialists.
    b. In all but two observations, both specialists, the subjects were first-term airmen of approximately equal experience and time in the service and measurable differences in performance probably did not exist.
5. The post-observation interview did not work as the maintenance people did not appear to discriminate task from environment.
6. The methodology was far too simplistic to observe, record, and interpret the richness and variety of both activities and

interactions in the environment with paper-and-pencil technology and one observer.

B. Results Pertinent to Maintenance--As the original methodology was applied to the research task the following distractions were noted.

1. Maintenance personnel spend a majority of their time waiting (for everything) and the slack time is seldom their fault.
2. Almost all maintenance is being accomplished by inexperienced maintenance personnel.
3. All things considered, the individual maintenance person does an excellent job in coping with distractions and eventually getting the job done.
4. The flight line maintenance environment of the POMO structure (AFR66-5) is substantially more turbulent than the AFR66-12 maintenance environment.
5. There is little technical assistance available from supervisory personnel.
6. Cross Utilization Training (CUT) is not working and there are people attempting to do maintenance on systems that they have only the vaguest idea about.
7. The crew chief's maintenance tasks are comparatively routine but the environment in which the tasks are accomplished is highly turbulent.
8. The decentralization of specialists into the AMUs means that that there is no central source of technical expertise.
9. The pressure to generate sorties is so intense and enduring that it overwhelms most other events in the environment.

Base Two. The basic methodology change was to increase the observation time to one full shift, approximately eight work hours, for each maintenance person observed. Along with the above change the post-observation interview was discontinued.

A. Results Pertinent to Methodology--As the modified methodology was applied to the research task the following issues were noted.

1. The increase of observation time to one full maintenance shift is definitely a step in the right direction. However, it is probably still insufficient. The amount of observation time required to acquire a representative sample of a maintenance person's activities and work patterns remains a research question and a question that is confounded by varying degrees of environmental turbulence.
2. The precise operational definition of "a task" remains a problem.
3. The comparison between a high performer's and a low performer's coping behaviors remains an unrealistic objective. It might be productive to determine average or normative behaviors and work patterns instead.
4. The availability of a hand-held computer into which observational data could be keypunched and which automatically recorded time would make this type of research much more productive.
5. Related to item #2, the operational definition of an environmental distraction is also a problem.

B. Results Pertinent to Maintenance--As the modified methodology was applied to the research task the following distractions were noted.

1. Maintenance persons are required to perform maintenance tasks for which there is no training, no technical data, and no test equipment.
2. Coordination when a supporting specialist or crew chief is needed is a continual problem.
3. First term airmen, with attendant skill levels, are performing almost all maintenance and generally without competent technical supervision.
4. The maintenance person works in an environment distinct and separate from the rest of a base's population.
5. In the specialists' environment events are driven by the random arrival of unscheduled maintenance demands.
6. Maintenance personnel function in an environment laden with uncertainty.
7. Transportation and communication are consistent major problems.

Base Three. The basic change to the methodology was to, if possible, increase the observer's unobtrusiveness. Interaction with the subject was minimized after introductions and while not rudely turning away interactions initiated by the subject, the observer did not encourage them.

A. Results Pertinent to Methodology--As the modified methodology was applied to the research task the following issues were noted.

1. The maintenance environment of each of the three major commands differs and the environments are different on each base, depending upon where you work.
2. Extreme weather conditions suggest that the magnitude and intensity of distractions do vary.
3. The observation time for any one subject is still inadequate.
4. The observation effort is spread over too wide an area of maintenance activities for this stage of research.

B. Results Pertinent to Maintenance--As the modified methodology was applied to the research task the following distractions were noted.

1. Severe winter weather obviously affects both man and machine.
2. The type and complexity of the assigned weapons system is a major component of the task environment.
3. Weapon system reliability and maintainability are increasingly important to the quantity of maintenance demands under extreme weather conditions.
4. On some tasks better grouping of tasks or scheduling could improve productivity.
5. Powered Aircraft Ground Equipment (AGE) is a continuing problem and its unreliability wastes large amounts of time.
6. Transportation and communication problems are compounded geometrically by extreme weather.

Discussion

Methodology. The methodology used in this study is too weak to provide hard data for policy making. The original intent of this study, however, has been satisfied because it was designed as exploratory research and to that end

it has been successful. While the study did not produce answers it did help to clarify questions, which was the anticipated outcome. The basic questions raised are: 1. operational definitions of tasks, distractions, and coping behaviors; 2. the complexity of the observation process, including sampling issues; and 3. the hypothesized existence of multiple maintenance environments rather than the assumed monolithic environment.

Maintenance. The predicted phenomena, and relationships within the phenomena, exist and the significance to increased productivity is higher than originally assumed. The three areas to be discussed here are the frequency and intensity of environmental (situational) distractions (constraints), the coping behaviors of maintenance people, and maintenance peoples' perceptions of what impedes productivity.

The frequency and intensity of environmental distractions is some function of the degree of environmental turbulence. The entire maintenance environment is more turbulent and less predictable than the clerical support environment and a significant portion of a maintenance person's time is spent dealing with environmental distractions. A substantial portion of the variance in the productivity equation can probably be accounted for by events external to the individual maintenance person and therefore beyond his or her control. The popular mythology is that the majority of the variance in the productivity equation is controlled by the individual maintenance person and the sketchy evidence from this study indicates that the situation, rather than the individual, contributes the majority of the variance. If future research confirms the preliminary data in this study managers could derive greater productivity increases at less cost by proper control of the environment as opposed to focusing all productivity efforts on the individual maintenance person.

The second observation deserving comment here is what it is that the maintenance person complains about. In most of the cases observed maintenance persons like their job but disliked environmental events surrounding it. They pointedly attacked segments of the maintenance environment and specifically perceived that, rather than assisting, environmental events hindered maintenance efforts. They are adamant and vocal about this, particularly since at two of the three bases no effort was made to collect this type of information. Not only do environmental events appear to control the majority of productivity variance from the observer's point of view, but the hands-on maintenance people verbalize a similar perception.

The third observation is that the individual maintenance person is doing a good to excellent job (sample not representative) and engages in reasonably positive coping behaviors from a productivity point of view. However, their coping behaviors are often contradictory to organizational policy, to them a perceived distraction. What they are experiencing is the frequently concealed conflict between pressure for increased productivity and compliance with organizational policy. The official point of view is that these two objectives are one and the same while the maintenance person views them as mutually exclusive. To the worker organizational policy and structure are part of the environmental distractions that steal valuable time from what is really important. Management might achieve significant productivity gains at low cost by better control of distracting environmental events. This strategy should produce quick returns and is easier to implement than significant attitude change.

AD-P003 306

LEADERSHIP INSTRUCTION AND OFFICER CANDIDATES
Some Unholy Thoughts

Commander John B. Washbush, PhD
University of Wisconsin-Madison

Abstract

Military people love the word "leadership." So often, however, when they use the term they emasculate it. It may come as a shock to many, but leadership is not an exclusive property of those in uniform! The problem is largely semantic as illustrated by the statement of the Navy Division Officer's Guide that "management" is something "leaders" do--most management literature sees it just the opposite. Careful attention to definition, a review of appropriate scholarly literature, and careful self-assessment can assist prospective officers in the important task of learning about leadership. Viewing "officership" as a combination of managership and leadership can help us cut a proper path through the folklore.

Introduction

My interest in leadership is rooted in my performance, a rather bad one as best I can recall, on a leadership theory question on my doctoral written comprehensive. I was bothered by the fact that I had taken a superficial attitude to the subject in my studies, and I was not really prepared to adequately address this important facet of organizational administration. Nonetheless, I passed the examination. However, I have carried on with a guilty conscience, and, as a result, have made the subject of leadership a focus of subsequent work in organizational studies. What I have discovered about leadership in the intervening years has left me in a state of frustrated confusion. I have come to the unholy conclusion that, to paraphrase Lord Russell, leadership is the science in which we don't know what we're talking about nor whether or not what we say is true. This dilemma is compounded because I am an officer in the military (which considers itself a paragon of leadership virtue) and am also charged with teaching the subject to officer candidates. My frustration is intensified because I am finally achieving some insight, but I do not find consistent scholarly support for my views nor am I very certain that I can communicate this view to students who are already "tainted" by the folklore. Let me outline my thinking on the subject and then give you a reprise of a recent teaching experience.

Leadership--Semantics and Myth

Some background may be useful. I am presently on a second active-duty tour in the Navy. Originally a history major, I completed the Navy's nuclear power training program after commissioning, served about 3 1/2 years of shipboard duty, and was a Naval ROTC instructor during my first active tour. I earned an MBA degree at the end of that period, and, after leaving active duty, continued my studies of administration in a doctoral program in educational administration. After completing these studies, I served as a faculty member in a college of business for 4 1/2 years. During all these years I not only had an opportunity to observe organizations in action but also to study them from a theoretical perspective. I have seen leadership defined from a variety of operational and theoretical perspectives. I have seen enough to be certain that the problem of semantics is the fundamental dilemma of leadership. There is also a problem of the "myth of leadership"--that military officers are leaders first and managers second and are therefore

fundamentally different from managers or administrators of all other organizations who are (poor souls), for some unexplained reason, managers first and leaders second. From an objective, organizational theory point of view, this myth is self-serving and confusing. Less politely, Rubbish!

Formal Organization, Management, and Leadership

If the logic of "leadership is whatever the boss says it is" can be applied--presumably we are all free to define the term to suit our own needs--I will take the liberty of casting my own pearls of wisdom for inspection. However, I must start back a bit in order to preserve a pretext of logic. Fundamentally, a formal organization is a group, structured in some manner, responsible for achieving some agreed-to purpose. Within these organizations, certain individuals are assigned positions of authority and responsibility (position power)--call them managers, officers, administrators, etc.--let me simply call them managers. In this vein, management is, in my opinion, authoritative decision making about the use of the resources of the organization for the purpose of achieving organizational goals. Note that I use the term "manager" generically, and, if you prefer, you may substitute "officer," "platoon leader," "commander," "quarterback," "wicked stepmother," or "SOB in charge." The point is that the people who fill these roles are functionally equivalent irrespective of whether or not a combat environment exists. Now I believe that leading is something these people NEED to do. Note, I said "need to do" not "do." I define leadership to be behavior, in a group context, that results in the willing compliance of the members. Note also that according to this definition anyone in the group can be the leader. The important thought is that those in positions of organizational responsibility can more easily and effectively meet their accountabilities if they are also effective leaders. Just as clearly, if leadership were to mean simply the behavior of those in charge (the "bad" definition of leader), then leadership would be anything they do, and we would have nothing to talk about! With my students, I insist on this behavioral definition for two reasons: (1) so that we can at least operate under the pretext of a common vocabulary, and (2) because I want them to feel that they are studying something that they need to understand and master in order to be effective when they get responsibility.

Managerial Behavior and Leadership

Let me develop a few more thoughts to round out this perspective. I believe that organizational situations must be met on their own terms, or, if you prefer, I believe that the logic of contingency theory must be given its due. Further, I believe that virtually all situations have both technical (task) and behavioral (relationship) elements present in varying degrees of importance. I believe that the person in position power will attempt to carry out the responsibility at hand by making a decision--e.g., to unilaterally determine the course of group action, to consult and order, to share in the group's deliberations, or to delegate authority and ratify action taken. Thus, I believe that there is a possible spectrum of behavior (note, I did not say "style"). Note, however, I am really calling this "managerial behavior," not leader behavior. I have found this decision-making approach to managerial behavior to be useful for a relatively simple reason. Given that I have defined what managers do (or what management consists in most basically), the question is, How? How means understanding the situation (facts, organizational policy and procedures applicable, and behavioral capabilities and patterns of the players--including one's self). The manager makes a decision on how to proceed and moves on (recall the decision spectrum I outlined above). Now sometimes the manager uses the authority (position-power) of the office to make and enforce a decision. Fine! But let's not call that leadership. It's power (or perhaps we might say "command"), and it's all right to call it that. What then of leadership? Recall my definition--behavior resulting in willing compliance. This has a persuasive connotation, and rightly so. But, let me add that this does not mean that I am a "blow-in-the-ear" person. Some people exercise leadership in a most bombastic manner (a la Patton and Lombardi), but that is both consistent with their personalities and effective. Others may use more subtle means with similar results. I believe that the individual is left with the dilemma of determining what works best. What a horrible thing to have to say

to a class of impressionable officer candidates! To admit that I, an O-5, a commander of the line, cannot give a simple answer to the leadership question is hard. To further imply that "That's your problem!" is even harder. But I believe that this is what we must do.

Leadership Theories and Their Failings

What then of so-called leadership theories? My perception is that most of these are misnamed. The popular theory group includes trait, leader behavior, Fiedler's contingency model, Vroom-Yetton's decision-making model, Path-Goal, Tannenbaum and Schmidt's leadership style continuum, Life-Cycle, McClelland's focus on power motivation, and the psychoanalytic postulations of Maccoby--among others. One common focus of all of this literature is the behavior of the person in charge. This focus is so intense that we are reinforced in the idea that leadership and being in charge are the same thing. However, the reality is that concern with the person in charge is the result of the need for effective management. In short, *managers can be more effective if they are also good leaders. In general, these theories describe* managerial behavior more than leadership behavior. Further, excepting Path-Goal theory, they ignore motivation in the respondents to leadership acts. Quite obviously the objective of any leadership attempt is to activate the motivation of others (Note, I did not say "motivate others" because that is psychologically impossible--motivation is interior and personal, and I may help you activate yourself but I do not activate you). Another failing of the present state of leadership theory is the lack of comprehensiveness. Finally, there is a general lack of insistence on self-evaluation on the part of the practitioner.

"Officership"--Management Plus Ten Percent

What then are we to do if we are to deal with semantic discontinuities and at the same time fulfill our responsibilities to officer candidates? We could start by clearing our minds of stereotypes. Courage and leadership are not the same, but military people love to romanticize on this point. Secondly, we need to remind ourselves that many of our own publications and manuals give at least lip service to the concept of persuasive leadership techniques. Thirdly, and most importantly, we need to stop the leadership vs. management debate. In this regard I propose the use of a new term, "officership." A friend of mine, a Navy captain who has had 4 commands (3 afloat) and is ABD in organizational psychology, has told me that, in his experience, being an officer is 90% management (taking care of business) and 10% leadership. That 10% is the difference between adequacy and excellence, and that is how he defines leadership. This is an excellent starting point. We really want excellent officership, and to achieve it we need excellent leadership. We should not be afraid to admit our own lack of sophistication in knowledge of the warp and the woof of this thing called leadership. We need to emphasize the responsibility of each prospective officer for developing skills in accurate self-assessment. In addition, we must commit ourselves to providing these people with the substance of leadership and motivational theories, imperfect though they are, thus giving them a basis for trial and assessment.

Socio-Technical Realism

I would like to expand a bit on this last point. In many services a great deal of attention is now paid to the technological sophistication of weapons and systems. There is understandably strong interest in expanding the technical sophistication of the officers who will supervise and use this modern, complex hardware. Accordingly, concern is well-placed for expanding the numbers of officers with technical education. But we must not forget that the services are fundamentally human institutions. Therefore, there is also an imperative to better understand the human side. Unfortunately, behavioral knowledge is sorely lacking in sophistication and precision compared to disciplines such as engineering and the physical sciences. These obvious failings, however, do not invalidate the social and behavioral sciences. Rather, I believe, these failings intensely dramatize the need for awareness and understanding.

This important challenge has two aspects that must be dealt with simultaneously. First, we need to commit ourselves to communicating to prospective officers a meaningful human behavior knowledge base. We could better do this if we had a substantial body of resources with stronger face validity than the general psychological, sociological, and business literature we so often have to use. For example, we need compendia of military case studies which focus on typical socio/technical problems and decision responsibilities junior officers face. Secondly, we must break down the barriers that we have ourselves constructed and which blind us to the obvious organizational reality that technology and human behavior are indivisible. We know that in environments where technology, resources, procedures and people inevitably interact, no decision is solely task or behavior-based. We must work to improve meaningful dialogue between technocrats and behaviorists.

Somewhat parenthetically, but relevant to this discussion, I would like to take one last swipe at those who lament the unpredictability of human behavior and seek to drown their sorrows with a swig of easy answers. The world, it seems is populated with vast numbers of latterday snake oil salesmen who feed on our insecurities and purport to know the answers to all the questions about deviant behaviors of people in organizations. The military services have grabbed the bait as vigorously as most other organizations. That is not all bad. After all, some of the salesmen are reputably trained and sell pretty good snake oil. Additionally, this is evidence that there exist commitments to making organizations both better places for people and more effective. But no one salesman has "the answer" and we cannot afford to try them all. Clearly the "deus ex machina" approach will not save us. Rather, we need to employ our substantial internal resources in making service people more comfortable and competent with the mysteries of human behavior. Clearly, we need to make our best efforts begin at the officer accession level.

Sifting and Winnowing

Earlier I promised that I would review for you a recent teaching experience I had with a class of NROTC seniors. Together we explored the mysteries of organization, management, motivation, leadership, and the challenges of being a junior officer. I tried to communicate and the students tried to understand. The result was something of a standoff. I am not sure that I ever succeeded in transferring the perceptions in my mind's eye. Clearly, that's an ego deflater. I also witnessed the power of images based on stereotypes and exacerbated by the semantic tangles we have invented for ourselves. However, I did achieve some success in forcing (note, I did NOT say leading) the students to confront a number of motivational and leadership theories and use them as a basis for (I hope) honest self-assessment. To the extent that I achieved that objective and also started them thinking about the responsibilities and challenges they will soon face on active duty, I feel some measure of success. In the meantime it's back to the drawing board to try to do a better job next time, mindful of the lessons learned. Come to think of it, that's not bad. In fact, that's what being an officer is all about!

Bibliographic Note

In To Get the Job Done, I have attempted to add some face validity to a body of organizational literature appropriate for use in leadership instruction of officer candidates. While much of the book is devoted to readings from the traditional sources of this literature, I made a definite effort to indicate its applicability to the naval service. The various section introductions in the text contain many "gems" of the same point of view expressed in this paper. The book is used in the senior year curricula of the Naval ROTC Program. The complete bibliographic citation is:

Washbush, John B. and Sherlock, Barbara J., eds.
To Get the Job Done: Readings in Leadership and Management, 2nd edition. Annapolis, MD: Naval Institute Press, 1981.

AD-P003 307

Windows On A Future World:
Human Dimensions of Air Force Doctrine in the 21st. Century

M. C. Devilbiss
Norwich University

Abstract

In 1983 the Air Force accepted the new Army combat doctrine "AirLand Battle" as a basis for joint tactical planning. This paper reviews this doctrine, examining the role of the Air Force and argues that in order for this new concept to be successful, a new kind of Air Force member with certain identifiable abilities and attitudes will be required in a battlefield of the future. The paper goes on to examine some of the implications of this hypothesis from the perspective of both the individual warrior and the combat team, and concludes by identifying some of the potential consequences to the Air Force at the organizational level in areas of accession and training of new personnel.

Introduction

In 1982, the U.S. Army's Training and Doctrine Command put forward a new combat doctrine, AirLand Battle 2000, designed to address itself to anticipated conditions on the battlefields of the twenty-first century. As is demonstrated by its name, conflicts on future battlefields are envisioned in this doctrine as taking place in at least two media, air and land, and close tactical cooperation in these arenas is emphasized. In October, 1983 this new Army doctrine was "accepted by the U.S. Air Force as a basis for joint tactical planning." (Army Times, October 10,1983: 19) The purpose of the present paper is to explore the AirLand Battle doctrine, focusing especially on some of its organizational implications for the U.S. Air Force, particularly in the areas of recruitment and training of personnel.

The Concept of AirLand Battle

An initial assumption of this doctrine is that future battlefields are quite likely to contain many high technology systems; sophisticated electronic and surveillance means, computers, lasers, and robots are expected to be present and extensively utilized. Secondly, battle "lines" on the battlefield of the future are envisioned as indistinct and in a constant state of flux. A third critical assumption of this plan is that although U.S. forces may face a superiorly armed and larger foe in the future, it is through maneuver, a technological edge, and highly trained personnel that U.S. forces can meet and overcome such a threat.

Contained within the doctrinal concepts are several operational implications. One is that inter-service (Army - Air Force) tactical coordination and planning is necessary to obtain strategic objectives. Secondly, this doctrine, as described by the Army itself, represents "a major 'evolutionary' change from past reliance on achieving the superiority in weapons and equipment..." "to a more intense, freewheeling style of warfare in which victory may hinge less on the firepower of mass formations and more on the abilities of small units to outmaneuver and confuse the enemy". (Wall Street Journal, January 22,1982:1) And finally, of particular importance here, the doctrine calls for decentralization of command authority and of decision-

making down to the company and platoon levels, and stresses the importance of small groups of combat teams who are capable of and skilled in rapid maneuver warfare. This is actually a return to the very old notion of the "elite unit" (see, for example, Thompson, 1982), but cast in a modern setting.

AirLand Battle: The Air Force Role

It is important to our purposes here to examine the anticipated role of the Air Force in this new tactical doctrine. Obviously, certain surveillance and intelligence gathering measures, close air support of ground forces, deep strike capability, and rapid air mobility of troops, equipment, and supplies - all elements of the AirLand Battle doctrine - are consistent with the Air Force mission and its strategic and tactical capabilities. But to achieve a more thorough understanding of the Air Force's role, it is necessary to view these macro-level goals from their basic building blocks at the micro-level of analysis. Here we can identify one not so obvious doctrinal congruence between the two services in the very notion of the combat team, the foundation upon which AirLand Battle is built.

The present author suggests here however that the Army's focus upon the combat team is at the secondary rather than the primary level of analysis. Rather, it is argued that the central concept in, and the key to the successful execution of, AirLand Battle is the individual military member, referred to here as the new "ALB warrior". This is not to say that the combat team is unimportant, for this is to commit the error in reverse: the fighter or bomber pilot or the missile launch officer is not the only individual involved in the successful execution of combat. The problem here lies in the matter of definition. With increasingly long-range weapons technology, does for instance even a fighter pilot have direct (i.e., personal) contact with an enemy - one that may only appear as an image on a screen or as a series of numbers representing a location? Likewise, those key support personnel (especially equipment maintainers), without whom the designated warrior would literally be unarmed, are part of the combat team, even though their "exposure" to the enemy may be even less direct than his is. What we are saying here is that the definition of the ALB "warrior" extends to all members of the team; new technology has made "combatants," in the broad sense, of more, not fewer, individuals. For the Air Force this means that the term, "the new ALB warrior," includes not only its rated officers, but personnel in certain non-rated areas (e.g., missiles) and key combat personnel in the enlisted ranks as well.

The crucial initial perspective then to adopt here is a social psychological one: to examine these implications within the context of individual expertise and development (primary level of analysis) and of small group behavior (secondary level of analysis) and then to relate these to organizational goals and means (tertiary level of analysis). An elaboration of this argument below shows how and why such an analysis is possible.

A New Individual for a New Doctrine

A major premise to be presented here is that a) the type of mission, b) the type of tactics, and c) the type of command structure specified in the new AirLand Battle doctrine necessitates the recruitment and training of a different type of warrior. With its missions emphases (hold or regain terrain; disorient, weaken, confuse, or destroy an enemy), its tactics (e.g., deep strike capability, rapid maneuver, ease of evasion, and

inability of detection), and its decentralized command structure, the battlefield of the future calls for a new kind of military member with certain characteristics. It is argued that not only are a different set of abilities functional to the Air Force of the twenty-first century, but that certain individual attitudinal components (previously de-emphasized at the expense of a concentration on abilities) are most important to identify. In turn, we shall examine the categories of abilities and attitudes and then go on to look at our secondary and tertiary levels of analyses.

Abilities. The "ALB warrior" will be an acknowledged expert in a particular area of military operations who will have as well an appreciation for and a working knowledge of the particular areas of expertise of other team members. The new ALB warrior will also possess a good general working knowledge of military operations as well as a knowledge of the language, customs, and history of the countries in which operations are being conducted. As such, the ALB warrior will be a knowledgeable expert, but also much more of a "generalist" than is now the case. This is necessary since the ALB warrior's role gives a wider latitude for action and requires ability to work not only with team members, but also with those of a larger unit or another service tasked with a similar mission, and to cooperate with officials and others in the host country. Thus the ALB warrior's training will be more inclusive in scope and therefore longer, getting away from the "assembly line" (unique specialization) approach. The ALB warrior's duties will also call for more responsibility for individual decision-making; therefore skill-based training will be incorporated with more general educational (abstract conceptual) development.

The ALB warrior must have the ability to plan as well as to act (rapid maneuver warfare and the element of surprise can place particular demands on these skills), and be able to think both defensively and offensively, as the new doctrine emphasizes both. Further, the ALB warrior must be fit physically as well as psychologically to withstand the demands of the environment, as well as frequent movement and stress.

Attitudes. Attitudes are less easily quantifiable than abilities, but are nonetheless often crucial to an individual's success or failure. Attitudes can be cultivated through exposure and reinforcement; they can also be possessed by an individual upon arrival to the military. It is suggested that the following (partial) listing of attitudinal components would be functional for the ALB warrior and should be encouraged by the military environment:

self-confidence and individualism - believes in own abilities; others recognize this individual as an expert

also a team player - recognizes and is unthreatened by the expertise of others; works individually and collectively to accomplish the mission

understands human relations - (especially important in this setting) - shows no prejudices; respects individual ability, the foundation of ALB

is dependable, consistently giving one's best efforts - internal motivation, not external threat or reward acts as an instigator of behavior; "getting by with the least effort" is unknown to the ALB warrior

possesses the notion of a social obligation to the state - sees the military in other than strictly vocational terms or as "just a job"

is willing to accept both risks and challenges; possesses both courage and integrity - the new ALB warrior is an individual not only of high ability,but of strong character as well

The Combat Team

The small combat team, the next level of analysis, is also an important element of AirLand Battle. As the development of human-machine interfacing becomes increasingly important with computer and robotic technology beginning to be used extensively on the battlefield, the ability to work not only with other people,but with programmable (and perhaps even environmentally responsive) machines will be a requirement. Furthermore, the intelligence and ability to repair the machine and/or to have a contingency plan in case of machine malfunction or failure will be a necessity.

If indeed individual ability is the foundation for the new doctrine, we may expect in the future to see "mixed" combat teams, since ability is no respector of race or sex. If the "contact hypothesis" holds, we may also then expect to see a decline in racism and sexism in the military, which would contribute to combat effectiveness.

We may also see at least some change in the concept of "leadership" where leadership is based upon expertise, as broadly defined. The present author does not suggest that the military will soon abandon rank as its traditional basis for authority, but rather that rank will become less important as knowledge and skill become more important, especially in less formal - decentralized operational - situations.

Organizational Implications

Increased emphasis will be placed upon the recruitment,training,and retention of the high quality ALB warrior. Selection programs may incorporate attitudinal assessment, as well as mental ability testing, into initial entry screenings; training programs will be longer and more inclusive in scope and will move toward the concept of <u>educating</u> as well as training the ALB warrior. A smaller size force is indicated, and given the high quality of individual necessary for this type of mission, salaries comparable to those of other professionals should be anticipated and planned for.

Recruitment must intensify and focus upon a small force of the very highest quality; lateral recruitment from those already at mid-career and increased recruitment of women and minorities is suggested. Because those recruited are increasingly likely to be married, a concept of rotational duty (similar to sea duty, with alternating blocks of time spent in and out of field operations) or of assigning couples together may prove beneficial. Because of the high quality of the individual ALB warrior, a potential problem is retention and replacement. To help address this, <u>individual</u> goals (e.g.,deepening or broadening one's level of individual expertise) must be identified and realized, if possible, through organizational opportunities.

Besides an increased need for technical skills, there will be a great need for human relations skills as well,and organizational behavior specialists, psychologists, and sociologists will become more important to the military. In fact, "human relations specialist" (as crucial as it

is to the small team concept) may eventually become a combat occupational specialty.

The emphasis in warfare will shift from physical destruction and loss of life to psychological crippling and debilitation of informational and logistical sources. The new ALB warrior will think this way and employ tactics designed to achieve these results. Moreover, because he/she has been educated as well as trained in military matters, the new ALB warrior will be able to think on both a tactical and a strategic level. As a result, we may eventually see a flowchart of advice and planning moving throughout the organization, rather than solely through the traditional "top-down" decision-making model of the past.

Relatedly, a very important caveat is in order here. The heterogeneity of the mission of the Air Force will not allow it to confine its future planning to the conventional warfare undertakings emphasized in the AirLand Battle doctrine. For instance, the concerns of the Air Force Space Command stress missions of scientific fact-finding as well as those of military intelligence or surveillance value. Those of the former may suggest a more cooperative, trans-national approach and philosophy while those of the latter suggest more nationalistic emphases. It is important to note that the ALB warrior, as envisioned above, is able to think and to act in both modes: while he/she is a loyal citizen of the nation, the ALB warrior is also able to recognize and deal with issues which may transcend national boundaries (e.g.,increasing scientific understanding of the earth, its weather and atmosphere, its natural resources, etc.). This is because the ability to think in global, as well as nationalistic, terms has been a part of this individual's leadership training in the new Air Force.

Finally, as has been suggested in the writings of several forward-looking thinkers, the military will have to create a new organizational identity and image of itself that is more compatible with a twenty-first century mission and values, and a new type of member in its midst. The Air Force, as the youngest service, is in the best position of all the armed forces to accomplish this. The influence of the individual on the military organization will be much more reciprocal than at any other point in its past,i.e.,the type of individual necessary will shape it to a great extent. In this respect then, the military may become more like other professions. In fact, a return to the model of the military as a profession is to be anticipated.

REFERENCES

Army Times, "AirLand Concept, NATO Tactics Differ", October 10,1983:19-20.

Thompson, Leroy, "A Short History of the Elite Forces" in Leroy Thompson and Michael Chappell, Uniforms of the Elite Forces,Blanderford Press, Dorset, England, 1982.

U.S. Department of the Army. AirLand Battle 2000. 1982 briefing, 20 pp.

Wall Street Journal, "Army Shifts Strategy to Give Smaller Units Room for Maneuvering," January 22, 1982: 1, 4.

AD-P003 308

PERFORMANCED BASED MANAGEMENT SYSTEMS: FILLING ARMY HUMAN RESOURCE DEMANDS IN THE FUTURE

Dr. Willem A. Hamel
U. S. Army Aviation Logistics School
Fort Eustis, Virginia

Abstract

Increasing qualitative and quantitative personnel requirements, shrinking entry level manpower pools, evolving battle doctrine, and advanced technology in increasing numbers of new systems with numerous organization changes are issues concerning the U.S. Army's ability to resource its force in the future. Because of the impact these issues have on how the enlisted force is resourced or is to be resourced, the Army is seeking ways within its system to insure that a "competency based force" is developed to meet the challenges of filling human resource demands in the late 1980's and through the 1990's. Within this context the Army is in the process of developing a new concept that will affect the way the enlisted force is structured, managed, and led. This concept, formally called "Force Competency", has been recently identified as the 'Performance Based Management System' (PBMS). Its goal is to increase the readiness of the force by retaining soldiers with appropriate combinations of leadership and technical competence.

The concept is a pilot study designed to evaluate the Military Occupation Structure (MOS) which recognizes soldier competence in two dimensions: leadership and technical skills. The notion is to attain competency of personnel to resource contemporary and future, sophisticated, high-tech equipment.

Since the American Army of the 1980's is faced with the dilemma of modernization in terms of fielding sophisticated technological equipment and initiating changes in organization structure, the Department of the Army Training and Doctine Command (TRADOC) identified the requirement (August 1982) to design, develop, implement, and evaluate a military occupational specialty structure which would recognize enlisted soldier competence in two dimensions: leadership and technical skills. This activity requires the recognition of the critical interrelationship between enlisted soldier structure, training, personnel management and Army unit mission. When finalized and implemented, the proposed structure is to provide for the measurement, recognition, and reward of individual soldier excellence at a mastery level of performance. In meeting this challenge of excellence, the Army realizes its need to increase current levels of leadership and technical skills competencies within an enlisted soldier career system. One that can be developed within existing Army authority. To the Army this means training and retaining competent personnel skilled in leadership and technical positions.

However, this may not in itself be sufficient, at least from an interservice perspective. Other military services will have to examine their peculiar enlisted personnel structures and undertake similar initiatives to enhance skilled personnel retention as well. Therefore, the major criterion for a model

occupational structure must include interservice adaptability between active and reserve military components, inclusive of a focusing on remuneration and benefits for skills that are in demand and/or in shortage. In addition, the accepted enlisted soldier military occupation specialty structure should facilitate lateral entry into the military service for personnel possessing advanced civilian skills having military application.

Army research indicates that the youthful population in the U. S. is in decline. Because of this the military will not be able to draw its normal entry level personnel from the traditional 17-22 year old population. And, by 1990, the now maturing Vietnam Era soldiers will be either egressing or retiring from active service, leaving the services dependent on untried combat forces. This condition will increasingly manifest, especially during the Army's modernization phase. Further, introduction of Robotics, high-tech equipment and structural unit changes will undoubtedly challenge personnel beyond present skill competencies. These inevitable changes signify a need for organization/personnel restructuring the attraction, retention and management of the American soldier in more complex environments.

Today's Army Force Modernization introduces complex job tasks which are direct manifestations of technologically complex, expanding and changing environments. The introduction of new tasks to meet technological criteria often displaces traditional tasks and roles of enlisted personnel. In the past, the Army simply assumed that personnel would be obtained and effectively trained to operate the equipment. The focus had been on hardware capability and not the human factors in the system. Consequently, present fielded systems have failed to achieve the desired effectiveness because soldier competency levels have been below standard during the systems developmental stages or were not adequately addressed. As a result, few junior noncommissioned officers (NCO's) are functioning effectively in leadership/supervisory roles; skill performance is questionable; and mid-level NCO's are unable to interface with the new technology. Because of the calibre of current soldiers the Army has been forced to reduce the number of complex tasks to be performed by soldiers. Further, the soldiers who may qualify for an occupational specialty under today's criteria, and who are trained in any given specialty, must perform specialty tasks, whether the tasks are for operating, maintaining or repairing the system. Unfortunately, contemporary enlisted personnel competency level is not where it should be. Therefore, this theoretical approach for transitioning to a more competent enlisted force is the Army's way of attempting to adequately address the skill requirements problem.

The overall goal is to improve the technical and leadership competency of the active enlisted force. This means providing a two dimensional success system that is performance oriented, can be certified for technical competence, accommodates changing soldier supply and requirements demand, provides compensation and status linked to level of performance, and is consistent with the Army Training 90 concept.

The Army Training 90 concept is concerned with training entry level soldiers only for their initial assignment. It is a unit level activity as opposed to structured skills training conducted in the various occupational schools.

A framework within which research and development could take place resulted from the combination of solicited proposals submitted by various Army schools.

Overall, leadership progression was viewed as occurring on a vertical axis (Exhibit 1), by leader level, based on existing measures to evaluate potential. Technical progression was conceived as occurring by competency and pay grade increase along the horizontal axis as determined by specified requirements of a particular skill and position grade using valid means to certify technical competency. And, combinations of technical and leadership requirements would, according to best hypothesis, provide vectoring along each of the axes.

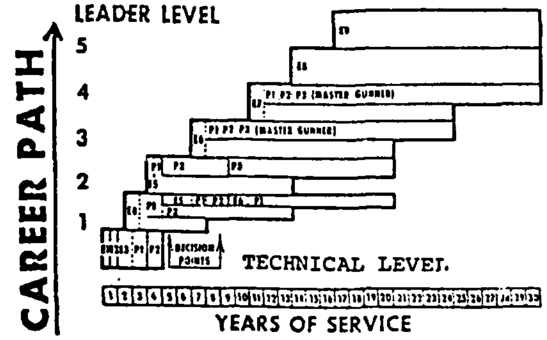


Exhibit 1

Decision points, i.e., points in time when determinations are made regarding the career mobility of personnel qualified for the next highest leadership position, were identified (Exhibits 1 and 2). In theory, enlisted personnel would be able to compete for available leadership oriented track positions when they reach the decision point stage. Similar opportunities would also be afforded to personnel who would select the technical career path (Exhibit 2) over the leadership track. Of course demand will determine the availability of positions. Hypothetically, enlisted personnel could be afforded "crossover points" from the leadership career track to the technical career path, or travel horizontally or vertically along both -- at least as far as acquired competencies and/or abilities would allow.

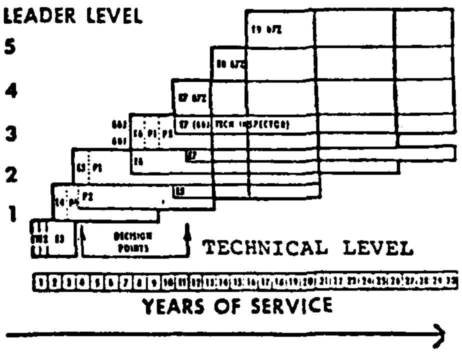


Exhibit 2

Theoretically, soldier personnel would progress through the ranks to Enlisted level 4 (E-4) status and choose to compete to progress horizontally along the technical dimension (Exhibit 2) at one point in their career. Later, if they feel qualified/competent, they may revisit their opportunity to progress further to Sergeant (E-5) at a higher level of leadership (Exhibit 1) and continue to track either way. These options have the built in advantage of providing continuous opportunities for extended careers. In addition, this concept, while retaining opportunities for leader development and attainment, would provide adequate compensation and status in lower levels of leadership responsibility.

As enlisted personnel progress in leadership attainment, technical skill proficiency acquisition, or a combination of both, competency in their duties or positions could theoretically transport them from apprenticeship positions along a given axis to master level positions of responsibility and job satisfaction. Currently, these superordinate leadership levels concentrate on leading and teaching subordinate personnel (Exhibit 3) while the increases in competency levels are oriented toward technical tasks.

In order to determine the feasibility of this concept, pilot programs have been initiated in five proponent installation schools of the U. S. Army: Armor, Signal, Military Intelligence, Surgeon General, and Transportation. The enlisted career paths of each pilot study are to highlight different approaches in providing common outcomes for consideration of different needs in each occupation specialty. These pilot programs also offer the best opportunity to generalize evaluation results to other occupation specialties and/or enlisted career management fields.

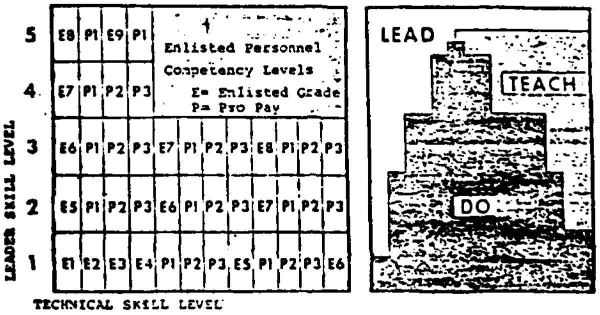


Exhibit 3

A logic system (Exhibit 4) was designed for this undertaking to insure that defined measures of merit would be consistent between pilot programs.

Each pilot program has specified means to certify technical competence and evaluate leadership potential at each leader and technical competency level. Enlisted personnel participating in these pilot programs are protected with a "stop-loss" management system to avoid potential adverse impacts on pay, promotion or school selection if the pilot programs are terminated.

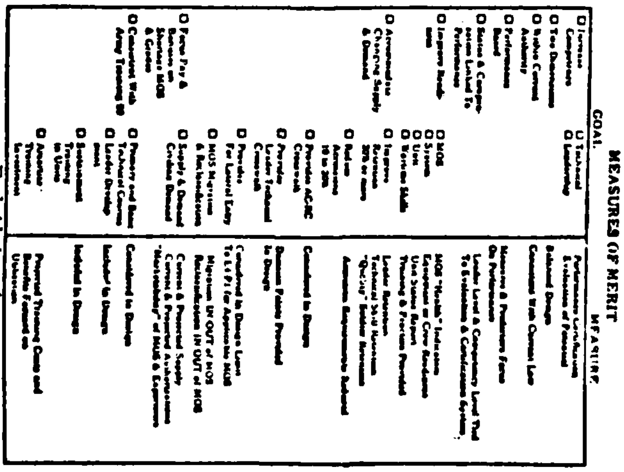


Exhibit 4

In summary, the former "Force Competency" concept now titled the Performance Based Management System (PBMS) will address the readiness of the force by retaining soldiers with appropriate combinations of leadership and technical competence. Under this system the 'up and out' enlisted soldier progression system is modified by a two dimensional success system. Soldiers are given the opportunity to earn more pay and achieve status for becoming more proficient within their technical areas of expertise as well as the traditional leadership progression track. This new system, based on performance, will exist along with the current system based on rank. This new PBMS approach will be available primarily for soldiers in Enlisted Personnel Management System skill levels two and three. Soldiers, therefore, will be offered the opportunity to earn pay increases for certified acquisition of technical skills within their occupational specialty, and concurrently have the option to compete for promotion in rank to senior leadership responsibilities or to retain lower level leadership positions. The Army is considering this change as a means of keeping highly skilled soldiers at team and section levels of leadership to better operate, maintain, repair and sustain systems; and to offer soldiers who are not interested in occupying senior leadership positions a chance to earn higher pay grades while staying on the job as a team or section crew leader, working directly with their equipment and junior soldiers.

REFERENCES

Force competency pilot program concept paper. U. S. Army Soldier Support Center, Fort Benjamin Harrison, IN., 1982

McDougal, Winn B. Force competency. Soldier Support Journal, 1983, 10, No. 1, 4-8

Wildrick III, Edward W. Soldier, what does industry want? Soldier Support Journal, 1983, 10, No. 1, 19-32

AD-P003 309

Computer-Automated Technological Innovation in Three Manufacturing Sectors

Ann Majchrzak
Krannert Graduate School
of Management
Purdue University

Veronica F. Nieva
Organizational and Marketing
Research
Westat, Inc.

ABSTRACT

A recent survey of a representative sample of manufacturing firms was undertaken to provide current information on the adoption of advanced technology and methods used for adapting workers to technological change. This paper describes results of the survey in terms of four issues: extent and nature of adoption of computer-automated equipment, extent and nature of industrial activities to train workers for adapting to technological change, relationships between automation and training, and organizational and environmental factors related to automation and training. Implications of the results for policymakers and students of technological innovation are discussed.

STUDY DESCRIPTION

As the productivity of American manufacturers has continued to fall behind that of Japan, West Germany, and other foreign countries, much recent attention and hope have been focused on technological advancements in our manufacturing industries. Two aspects of technological advancement that have been of major interest to policymakers is the adoption of advanced technology and methods for adapting workers to technological change (e.g., the Spring, 1983 congressional hearing on manufacturing automation). Policymakers have been intent on learning the extent and nature of new technology that currently exists, as well as the extent and nature of existing industrial activities for preparing workers to adapt to the new technologies. In addition, there has been an increasing concern for identifying factors that affect these aspects of technological advancement so that supportive governmental initiatives can be adopted (Tornatzky et al, 1982).

Given these concerns, a survey was sponsored by the Congressional Office of Technology Assessment in 1982.[1] The survey gathered information on the extent and nature of computer-automated equipment in use (including robotics, CNC, DNC, materials-handling, CAD, and computer-automated storage and retrieval), and the extent and nature of education and training for new technology in place. Respondents were representatives of 303 plants in the three manufacturing industries most likely to automate (Gunn, 1982): transportation equipment, metalworking, and electric and electronic equipment. The plants were selected by means of a multistage, stratified probability sampling procedure in order to represent all United States plants in the three industries. The survey was conducted as a telephone interview with respondents carefully screened to ensure

[1] The survey was conducted under the direction of the authors under OTA Contract #233-61700.

knowledge of the subject matter. In addition to information on adoption and adaptation, selected organizational variables such as age and size were also gathered. Finally, information was obtained on environmental variables, such as growth in industry employment, wage, and productivity were obtained for each four-digit SIC code included in the sample.

Analytic results of the survey will be discussed in terms of four issues.[2] First, analysis of the extent and nature of adoption of computer-automated equipment in the three industries will be described. In this analysis, updated information on the prevalence of technological innovation in the manufacturing sector will be presented. Although there has been substantial concern about technological innovation, there has been little current empirical evidence documenting its prevalence. Information that does exist tends to be either outdated (e.g., GAO, 1976) or narrowly focused on one technology (e.g., BLS, 1982). In this study, the extent to which a range of technologically advanced equipment has been adopted will be described.

In addition to a discussion of the extent of adoption, the current nature of computer-automated equipment use will be described. Much of the research to date on technological adoption has been constrained to describing whether or not certain technologies are in place. In this study, analysis will go beyond measuring extent, to describing patterns of technological adoption. In describing patterns, the concurrent use of types of technological equipment as well as the extent to which this equipment is integrated with other equipment on the shopfloor will be presented. For example, such questions as the following will be addressed: Do plants which use robots also tend to use CNC and not CAD? Are certain types of automation use more likely to be related to the integration of shopfloor equipment than other types? In addition, automation profiles of plants will be developed. The profiles will distinguish plants with similar patterns of technological use.

A second issue to be addressed by the analysis will be the extent and nature of industrial activities to train workers for adapting to technological change. Although there has been substantial recent research on expected changes of workers as a result of technological change (e.g., Rumberger, 1981; Riche, 1982; BLS, 1982; Duke and Brand, 1982; Walton and Vittori, 1983), none of this work has systematically addressed the activities actually being undertaken by the firms. In this analysis, the extent to which technologically innovative manufacturing plants are currently sponsoring education and training programs for workers is described. In addition, the nature of this training will be specified by examining relationships among several characteristics: percent of workforce receiving training, occupational groups receiving training, skill and knowledge areas covered, format of training, and company policies about training outside the firm. In this analysis, such questions as the following will be addressed: Which skills and occupations are receiving the greatest attention by firms concerned with technological adaptation? Does the amount and nature of skills training vary with the occupations to which it is addressed? In addition, training profiles of plants will be developed. The profiles will distinguish plants with similar patterns of adapting workers to technological change.

---

[2] Since results of the survey are not publicly available until January 30, 1984; actual findings cannot be discussed at this time.

The third issue in the analysis will focus on the relationships between automation and training. In this analysis, the nature of training activities most likely to be provided with certain technological advancements will be identified. For example, such questions as the following will be addressed: Is the use of CAD more likely than other technologies to be coupled with skills training for production engineers and programmers? Are plants that use a diverse group of technologies more likely to provide training in a range of skills (e.g., basic physical science, safety, etc.) than plants that use only one type of equipment?

The final issue to be addressed by the analysis will concern the interrelationships between the organizational, environmental, automation, and training variables included in the study. Although there has been considerable research or factors affecting technological adoption, the research has either been focused exclusively on aggregate-level variables or based on case studies of single adoption decisions (Tornatzky et al, 1983). Furthermore, much of the research has focused primarily on factors related to the extent of adoption, providing little insight to factors related to patterns of adoption or the extent and nature of adaptation. Therefore, in this analysis, the interrelationships between organizational, environmental, automation, and training variables will be described. Such questions as the following will be addressed in this analysis: Is the past wage growth in an industry related to the diversity or type of technological equipment adopted? Are recent changes in an industry's employment level related to the extent or nature of training provided by firms in the industry? Does an organization's size mediate the relationship between automation and training? Based on this analysis, a preliminary model of factors related to adoption and adaptation will be presented. The comments of industry experts concerning the adequacy of the model for describing their experiences will be shared.

The results of the analysis on these four issues will raise several implications for policymakers and students of technological innovation. The paper will close with a discussion of these implications.

REFERENCES

Bureau of Labor Statistics, Technology and Labor in Four Industries, Washington, D.C.: U.S. Department of Labor, January, 1982.

Congress of the United States. Hearing of the Joint Economic Committee, Subcommittee on Economic Goals and Intergovernmental Policy. "Manufacturing Automation: Selected Employment and Education Issues," March 18, 1983.

Duke, J. and Brand, H. "Cyclical behavior of productivity in the machine tool industry," Monthly Labor Review, 1981 (November), 27-34.

Gunn, T.G. "The mechanization of design and manufacturing," Scientific American, 1982 (September), 247(3), 114-130.

Riche, R.W. "Impact of new electronic technology," Monthly Labor Review, 1982 (March), 37-39.

Rumberger, R.W. "The changing skill requirements of jobs in the U.S. economy," Industrial and Labor Relations Review, 1981, 34(4), 578-590.

Tornatzky, L.G. et al, The Process of Technological Innovation: Reviewing the Literature, Washington, D.C.: National Science Foundation, May, 1983.

U.S. General Accounting Office, Report to Congress: Manufacturing Technology; A Changing Challenge to Improved Productivity, Washington, D.C., 1976.

Walton, R.E. and Vittori, W. "New information technology: Organizational problem or opportunity," Office: Technology and People, 1983, 1, 249-273.

AD-P003 310

PROTOTYPE DEVELOPMENT OF AN INFORMATION-SHARING AND DECISION SUPPORT SYSTEM FOR THE MANPOWER PERSONNEL AND TRAINING COMMUNITY

Joel Oxman, Lois Richards, and Linda Loughnane
MATRIS Office, San Diego, CA

ABSTRACT

The availability of up-to-date and reliable data on the substance and funding of Research and Development efforts within the Manpower Personnel and Training (MPT) community, and the valid linkage of such data through a systems approach, are important aids to researchers and managers in the Department of Defense. The Manpower and Training Research Information System (MATRIS) is a computerized, information-sharing and Decision Support System (DSS) designed to provide such aids to those involved with the conduct and/or fiscal management of Department-of-Defense-sponsored, people-related Research and Development pursuits. Although already in operation, the evolution of MATRIS continues within the framework of the prototype development model. The prototypal development process of MATRIS, the structure and content of its data base, and the services and products which the system makes possible, are described.

Introduction

The pursuit of Research and Development in the Manpower Personnel and Training (MPT) area has, over the past four decades, resulted in significant gains in training technology and manpower and personnel management procedures. Accompanying the substantial progress made in these regards has been an "information explosion" of unprecedented size. The proliferation of technical journals, Research and Development management literature, and conferences in the MPT area attests to the ever-increasing need for, and challenges of, the efficient and timely sharing of information. It is with this issue -- of the information and planning requirements and resources of MPT researchers, planners and policy-makers--that the present paper is concerned. The challenges and potentials of a reliable information-sharing network in the MPT area will be addressed in this paper, and discussed in regard to the development of a prototype, computerized system designed to serve both as an information-sharing system and as a Decision Support System (DSS).

MATRIS -- the Manpower and Training Research Information System -- stores, manipulates and retrieves data related to the substance and funding of MPT Research and Development efforts throughout the U.S.* The system provides service to researchers, managers, planners and policy-makers concerned with the current status and the future directions of people-related research which is sponsored by the Department of Defense (DoD). The primary aims of the MATRIS program have been to develop and evaluate a prototype system which can: (1) facilitate the sharing of reliable and valid information among researchers in the MPT community, and (2) assist managers, planners and policy-makers at all DoD levels with the optimal development and utilization of Research and Development programs.

Prototype Development

The MATRIS program began with the implementation of two, independent information systems -- the Training and Personnel Systems Technology (TPST) and the Research and Development Information System (RDIS) -- each designed to track unique but limited aspects of ongoing Research and Development efforts. TPST

and RDIS served as feasibility projects, with emphases on the delineation of useful data elements, the identification of necessary data sources, and the initial formatting of the data bases. With the resolution of these primary issues, TPST and RDIS were shown to be viable information systems. Subsequently, MATRIS's objective became that of merging and refining the TPST and RDIS systems into one comprehensive DSS which could provide potential users with up-to-date information on the substance and funding of MPT research and development efforts. Importantly, MATRIS was envisaged as an integrated information network with the capability of tracking and linking MPT research and development efforts across various levels (e.g., congressional category, budget category, program element, project, work unit) of the DoD hierarchy.

In concert with Revell's (1981) concept of organizational schema, the MATRIS program was staffed and developed with a continual concern for the environmental context of the DSS. That is, high priority was placed upon the responsiveness of MATRIS to the information requirements, organization characteristics and work objectives of immediate and potential users of the system. Both the organizational (Moynihan, 1982) and the informational (Solotruk & Kristofic, 1980) integration of the MATRIS program were seen as high priorities. Evaluations of MATRIS's operational capabilities over the two years of its prototypal history have indicated the reliable storage and valid linkage of data in the system's 200 data element fields, and have served to document the system's momentum in achieving its goals of information-sharing and decision support within the MPT community. As a DSS embedded within a complex organizational framework and concerned with the substance and planning of rapidly changing Research and Development efforts, it is the aim of MATRIS to remain flexible both in the technological sense (Solotruk & Kristofic, 1980) and in regard to the dictates of effective management techniques (Simon, 1977).

System Configuration

MATRIS storage and retrieval functions are carried out on a Univac 1100/60 computer located at the Defense Technical Information Center in Cameron Station, Alexandria, VA. The updating and data processing functions (e.g., the abstracting and indexing of textual material, entry of current fiscal data) of MATRIS are performed on a Harris/100 computer, located at the Navy Personnel Research and Development Center, San Diego, CA. The BASIS Data Management System (Batelle, 1981) is employed for data storage, manipulation and retrieval functions. MATRIS terminals are in operation at both the San Diego and Alexandria locations, with direct system access available and with information transfer between the Harris and Univac computers accomplished via telephone communication lines. Hard copy products can be generated and received at both the San Diego and Alexandria operating stations.

Structure and Content of the Data Base

The structure of the MATRIS data base is hierarchical in nature, and reflects the general organizational structure of Research and Development within the MPT community. MATRIS is concerned with data and information pertaining to three different funding levels: the program element, project and work unit levels.

Program elements relate to the four-dimensional categorization of MPT research according to: (1) DoD funding/strategic priorities (MATRIS tracks only "Research and Development" program efforts.); (2) the general nature of the Research and Development effort (i.e., basic research, exploratory development, etc.); (3) equipment/activity type (e.g., military sciences, aircraft and related equipment, etc.); and (4) branch of service. The project level is

concerned with major Research and Development "thrusts", "objectives" or "goals" (e.g., Simulator Development, Operational Flight Training) which are derived from the larger concerns defined by the various program elements. Information and decision support aids pertinent to these first two levels are used primarily by Research and Development managers, planners and policy-makers. The work unit level, which is typically of primary concern to researchers and developers in the field, focuses upon data and provides for information related to the specifics (e.g., methodology, research findings) of past, current and planned Research and Development efforts.

Each of these three levels includes data pertaining to fiscal concerns as well as information regarding the nature and objectives of the work being conducted, the organizations and individuals involved, and data element links which allow for the examination of relationships between the various organizational/funding levels. The three levels of funding data and research information are subsumed in hierarchical fashion by the four Congressional Categories (Human Factors, Manpower and Personnel, Education and Training, and Simulation and Training Devices) which serve to partition the major goals and objectives of MPT Research and Development.

Services and Products

To the extent that the MATRIS program provides the right information to the right people, at the right time, and in a cost-effective manner, it may be said to fulfill its functions as a DSS (Mader & Hagin, 1974). The accomplishment of these objectives depends not only upon the hardware and software configurations of the system, but also upon the continual processes of data collection and the valid reduction and reliable coding, indexing and updating of information pertinent to the MPT community.

The major products of the MATRIS program are:

(1) The Training and Personnel Systems Technology Research and Development Program Description (the "Budget Book") -- a document which provides an overview of the MPT program. It is published annually, though updated segments can also be generated at any point in time. It contains Program Element and Project synopses, and fiscal information derived from the President's Budget. Laboratory planners and managers, and service and OSD (Office of the Secretary of Defense) headquarters personnel use this document as a forecasting and decision support device.

(2) The Directory of Researchers -- a document which lists those individuals who are performing and/or managing current, people-related Research and Development for the DoD. It is published annually, but also can be updated and produced in segments upon request. The Directory has served both researchers and managers as a useful aid in their pursuit of open communication lines within the MPT community.

(3) Subject matter retrievals -- information provided through systematic and structured searches of the data base along lines prescribed by interested researchers. Such retrievals are based upon the reliable, subject matter indexing of textual data at the work unit level. MATRIS staff index each work unit in the system through the use of the Human Resources Research Indexing Vocabulary (HRRIV) -- a unique, hierarchical, concept-based indexing system developed for MATRIS to capture the essense of MPT research efforts. At present, the HRRIV consists of approximately

1800 vocabulary terms. It is a dynamic indexing system which is continually exposed to the processes of conceptual refinement, term additions and appropriate hierarchical reorganizations.
(4) Program element/project listings -- information on the funding bases and fiscal plans for selected program elements and/or projects of interest. This type of product has been useful to program managers and planners in their review of past DoD fiscal priorities and in their anticipation of future budgetary trends and goals within the MPT community.
(5) Multi-service retrievals -- documents which display fiscal and/or research content information (e.g., program goals and priorities) across the various DoD services. This type of information may be useful to planners and managers in avoiding duplication of effort across the service branches, in keeping up-to-date on changing Research and Development priorities within the MPT area, and in monitoring the course of progress and budgetary expenditures within the various service branches and program element/project categories.

Also available as products from the MATRIS program is a myriad of documents based upon specifically framed inquiries regarding the fiscal details and/or content of Research and Development efforts within the MPT community. The initiative and inquisitiveness of MATRIS users, and the systems and research knowledge of the MATRIS staff, set the boundaries for the present and potential utility of MATRIS as a DSS. As is true in any area of technology or information management (Ackoff, 1981), it is the human component in the MATRIS system which is ultimately responsible for the optimal realization of the system's potential.

Future Directions for MATRIS as a DSS

To date, MATRIS has developed and functioned within the context of a Research and Development atmosphere, progressing in an iterative fashion through development, evaluation and refinement stages. As MATRIS continues, as an operational system, to be responsive to the needs of those within the MPT community, this prototype development process of development-evaluation-refinement remains as a necessary and vital aspect of the system. The MATRIS staff continues to pursue a number of issues which are central to the system's optimum use and efficacy.

One current area of high priority for instance, concerns the further determination of user information and decision support needs. A good deal of staff resources are presently being devoted to more clearly defining the course of information flow and the nature of decision making processes (re. Research and Development priorities, funding plans and allocations) within the MPT community. In parallel with addressing this issue, of course, there is also concern for identifying those additional data sources which will be required to provide the appropriate input to the MATRIS data base. MATRIS's future plans include the promotion of more information-sharing among the various DoD laboratories and between service branches, and the concomitant provision of a "secure" data base (or distributed data bases) with varying levels of privacy and user access.

Accompanying such plans for refinement and expansion of the MATRIS data base are subprojects dealing with both software and hardware innovations and acquisition. The construction and employment of expert systems, the use of artificial intelligence techniques, and the integration of political, social, etc.,

factors into the DSS framework are just a few examples of the technological challenges and opportunities which exist (Feigenbaum & McCorduck, 1983).

Other areas of concern include: the implementation of graphics capabilities to supplement user products; the entry and updating of specified data via remote terminal sites; the continual refinement and restructuring of the HRRIV -- MATRIS's indexing system -- to accurately and specifically capture the character of Research and Development efforts being pursued; and the extension of the prototype system designed by the MATRIS program to other technology areas.

MATRIS was developed as a prototypal system in response to the needs of Research and Development personnel concerned with the direction and future of DoD-sponsored, people-related research. As a now fully operational system, MATRIS continues, through the encouragement of information-sharing and the provision and revision of decision support aids, in its developmental process towards the goal of serving the MPT community.

References

Ackoff, R. L. Creating the Corporate Fututre. New York: John Wiley & Sons, 1981.

BASIS Data Management System. Columbus, Ohio: Batelle Development Corporation, 1981.

Feigenbaum, E. A., & Mc Corduck, P. The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World. Menlo Park, CA: Adison-Wesley Pub,. 1983.

Mader, C., & Hagin, R. Information Systems: Technology, Economics, Applications. Chicago: Science Research Associates, Inc., 1974.

Moynihan, T. Information systems as aids to achieving organizational integration. Information & Management, 1982, 5, 225-229.

Revell, N. Managing the development of data base systems. Inforation & Management, 1981, 4, 197-205.

Simon, H. A. The New Science of Management Decisions. New Jersey: Prentice-Hall, 1977.

Solotruk, M., & Kristofic, M. Increasing the degree of information system integration and developing an integrated information system. Information & Management., 1980, 5, 207-212.

* MATRIS is an activity of the Defense Technical Information Center (DTIC), and is accessible only to DTIC-registered users. For further information, contact Lois Richards, Chief, MATRIS Office, DTIC-R, San Diego, CA 92152 (619-225-2056; AV933-2056).

AD-P003 311

An Identification of Actual vs. Reported Accident Causes

Dr. Harv Gregoire
Institute of Safety & Systems Management
University of Southern California

Abstract

This initial research developed a survey technique designed to uncover specific areas of weakness in organizational safety. It was postulated that data concerning insurance claims, medical treatment, and other recorded accident cause variables may be contaminated since certain accident causes are either not reported or erroneously reported for a variety of issues such as fear of blame, misperceptions, and other intentional and unintentional reasons.

Introduction

Background

The research reported herein describes an attempt at developing an approach to uncover actual causes of accidents or near accidents in which causes may not have been reported accurately or not reported at all. The fact that many insurance forms and accident reporting forms are perceived as efforts to fix blame or assign culpability in accidents may bias some of the very data that they are designed to derive. While there is an undeniable need to use the multitude of such forms which exist for individual accidents, there is a separate issue of identifying generic causes of accidents in organizations throughout industry and government.

It has been postulated that in questionnaire surveys, respondent anonymity will result in more objective data gathering than those surveys which require respondent identification. The reasons are usually found in the hidden corners of that nebulous area known as human nature. In responding to a questionnaire, most people do not want to appear significantly "different" from fellow respondents. Most importantly, respondents generally do not want to appear any less careful, responsible, or worthwhile than contemporaries whether the roles involved are that of employee, associate, or insurance client. It may be a combination of social pressure and human pride that often causes respondents to answer questions in terms of what they perceive the interrogators want as data or what is socially acceptable rather than what may be objective truth as they know it. We will probably never discern how much questionnaire data has been biased in the desire to answer questions in a socially appropriate manner rather than a completely objective manner.

Method

A trial Accident Cause Evaluation (ACE) consisting of an interview and questionnaire protocol was developed to access the issue of incorrectly reported accidents. Variations of the main questionnaire were administered in an iterative fashion to ten groups of high school and college students in order to modify and improve wording and format until it was decided that maximum clarity had been established in the ACE questionnaire. As a pilot study, a small fifty to sixty person military vehicle maintenance organization was evaluated on an experimental basis to determine the utility and practicality of such a device. The prototype ACE device was administered to the entire unit, including supervisors, maintenance workers, and administrators. There were forty respondents in the trial research.

Discussion

Portions of the trial research questionnaire and resultant data are reproduced and discussed below. It should be emphasized that the author's post-item comments pertain only to this trial sample administration and are not meant to be generalized across organizations.

Accident Cause Evaluation Questionnaire

In each question, the term "the accident" refers to an accident or near accident at work you are describing which was incorrectly reported as to its cause. (Note: Answers stated in percentage of respondents)

1. Rate the accident on a scale from 1 to 10 in terms of seriousness (1 = minor, 10 = major, or N/A = "close call"). - 10 = 53%; 7 = 13%; 3 = 25%; N/A = 8%. Over 50% recalled serious accidents in which causes were misreported.

2. Describe the accident in brief terms: - Various descriptions.

3. Rate how certain you are on a scale from 1 to 10 that you know what caused the accident (1 = uncertain, 10 = positive). - 10 = 80%; 8 = 10%; 1 = 10%.

4. Rate how certain you are on a scale from 1 to 10 that the item reported as causing the accident did not cause it. (1 = uncertain, 10 = positive) 10 = 80%; 8 = 10%; 1 = 10%. Both questions 4 and 5 indicate a high reported certainty of actual versus reported casses.

5. How long ago did this incident occur? A = within the last month - 13%; B = within the last 2-6 months - 43%; C = 6 months to a year ago - 25%; D = over a year ago - 20%.

6. What time of day did the accident occur? A = 1:00am to 5:59am - 10%; B = 6:00am to 11:59am - 15%; C = 12:00pm to 5:59pm - 43%; D = 6:00pm to 12:59pm - 32%. As reported, pm was more hazardous than am.

7. The condition of the person or persons involved in the accident was A = well rested, alert - 25%; B = fatigued - 13%; C = overworked - 10%; D = emotionally distracted - 12%; E = bored - 10%; F = sleepy - 20%; G = unknown - 10%. Less than 25% implicated high workload as causative.

8. In your opinion, the accident was: A = preventable - 63%; B = not preventable - 20%; C = uncertain - 7%; D = not applicable - 10%.

9. If this incident was merely a "close call" and not an actual accident, was it reported? A = yes - 3%; B = no - 15%; C = N/A - 82%. A significant majority of respondents to this question indicated an awareness of incorrectly reported accident causes. Herein may be the value of this instrument. If this sample is representative, there appears to be a definite need to identify actual vs. reported accident causes if those accidents are to be prevented.

10. Which of the following categories of causes best describes the reported cause of accident? A = material failure - 40%; B = procedural error - 5%; C = communication problem - 10%; D = equipment design - 15%; E = inadequate training - 5%; I = negligence - 5%; J = fatigue - 5%; K = unknown cause - 10%; L = other (if L is selected, describe briefly) - 5%. Compare the two most frequently reported causes, material failure and equipment design, with those causes reported as actually causing the accident in question 11.

11. Which category in question 10 (above), actually caused the accident? (If L is selected, please describe briefly.) If you list more than one answer, please list them in order of importance. A = 10%; B = 30%; C = 13%; D = 10%; E = 5%; I = 25%; J = 13%; K = 10%; L = 5%. These answers totaled over 100% of the respondents since many insisted on identifying two "actual" categories of cause (procedural error and negligence). In subsequent interviews, it was determined that attempts at preventing these two causes would be more effective and less expensive than action formerly planned by management.

12. In your opinion, which of the following answers describes the reason why the cause of the accident was not accurately reported. If you list more than one answer, please list them in order of importance. A = person reporting cause didn't know - 5%; B = embarassment - 35%; C = wrong cause reported due to incorrect or missing data - 5%; D = fear of blame and consequences - 30%; E = insurance would not pay damages or medical compensation - 15%; F = reporting the actual cause wouldn't result in any improvement - 0%; G = wrong cause or no cause reported to avoid involvement in accident investigation - 10%; H = other (describe)- 0%. Inaccurate reporting was identified as being caused primarily due to embarassment and fear of consequences due to blame. It appears that the vast majority of cases where incorrect cause reporting occured was intentional. If this trial research had been limited to organization - specific accidents, one could speculate that a major safety challenge to management would be to devise an acceptable, no-fault, accurate

reporting procedure in order to identify and prevent actual accident causes. This issue of "no-fault" accurate reporting is the subject of future research.

13. What could be done to prevent what actually caused the accident or "close call" from happening again? Among the wide variety of "prescriptions" offered, the most common theme concerned "following proper procedures" or "not taking (procedural) short-cuts".

14. Would you have answered this questionnaire exactly as you have answered it if it would not have been anonymous? A = yes - 33%; B = no - 55%; "not sure" or "depends" was listed as an answer by 12% of the respondents. Anonymity does appear to elicit more objectivity than respondent identification. However, during data analysis and subsequent interviews, it became obvious that this question needed (A) modification, or (B) other data to make results more interpretable. A researcher would need to know if the individual data per questionnaire pertained to a personally experienced accident or to an accident which occured to someone else. Since fear of blame was such a major issue, the results of this question would probably be affected by personal involvement. 12% "not sure" or "depends" response to this item probably indicates this personal involvement concern.

Thank you for your time, your identity will remain anonymous. Your time, effort, and integrity in answering this questionnaire is a significant contribution to increased safety.

Conclusion

A questionnaire and interview protocol was administered to a select population sample to identify what erroneously or non-reported variables have been implicated in causing major and minor accidents in the experiential history of the respondents. Since anonymity was designed into the survey and guaranteed to the respondents, improved accuracy of variable identification resulted. In the pilot study, an analysis of variables cited was conducted in order to identify which categories of causes were reported as actually versus reportedly responsible for accidents in the particular organization sampled. Data from the organization sampled identified procedures which should be revised versus the procurement of new equipment as a measure to be taken in reducing accidents.

The subject questionnaire is being revised and hopefully improved since the trial research. Findings thus far indicate probable low utility or minimal value in using this approach for any specific accident investigation. There does appear significant potential however for improved organization - wide generic accident prevention if management had the means to gather the data indicating that there were significant differences in causal versus reported variables implicated in accidents, and if causes could be accurately identified. A more effective use of resources could be allocated to alter the pertinent conditions, people, procedures, or equipment.

AD-P003 312

Relationships of Biorhythms and Categories of Industrial Accidents on an Air Force Installation

Carl B. Freitag and Connie M. Haefele
Department of Psychology, Middle Tennessee State University
Murfreesboro, Tn 37132

Biorhythms were calculated for 852 personnel involved in industrial accidents, over a six year period, on an Air Force installation. Each subject had biorhythms calculated and plotted, by computer, for the month of their accident and were analyzed for relationships between critical days and negative phases of each cycle. Chi Square analysis indicated no significant relationships for the entire group. The catagories of accidents called "Unsafe Personal Factors" and "Unsafe Practices" were individually studied as well as the specific causes of "Work Posture/Position," and "Inattentive, Absent-minded." These were felt to be most likely to be associated with biorhythm cycles. No accident frequency was found beyond the expected level. A unique feature of this study was to generate new random birthdays for all subjects. The same procedures were followed using the new biorhythms plots for the month of their accidents. Results showed the same insignificant relationship, and it was concluded that there was no evidence to link biorhythm cycles and accidents.

Introduction

There has always been much popular interest in systems and approaches to predicting future events. Recently, biorhythm theory has been added. One can now get newsletters, books, and calculators to determine your daily biorhythm phases. This is a theory that postulates that people function on physical, emotional, and intellectual cycles of 23, 28, and 33 days respectively. These sine wave cycles have positive and negative phases as well as critical phase days where the curve changes from positive to negative and vice versa. The theory states that the cycles start at birth and continue throughout an individual's lifespan, and an individual's performance will be affected by the place one happens to be on a cycle. During the physical cycle, one's physical strength, endurance, energy resistance, and confidence will be affected. The emotional cycle will affect one's sensibility, nerves, feelings, intuition, cheerfulness, moodiness, and creative ability. The intellectual cycle will affect intelligence, memory, mental alertness, logic, reasoning power, reaction, and ambition (Thommen, 1969). Supposedly, when any curve crosses the ordinate, this makes it a critical day, and susceptability to accidents increases. Of course, this has gained the interest of industrial safety offices which would greatly desire this predictive power. The research in this area, however, is mixed. Evidence for support of the theory has been reported (Anderson, 1973; Gittelson, 1975; Pittner and Owens, 1975; Schwarz, 1976; Thumann, 1977). Studies demonstrating lack of support include , (Khalil & Kurucz, 1977; Shaffer, Schmidt, Zolotwitz & Fisher, 1978; Wolcott, McMeekin, Burgin & Yanowitch, 1977).

Chaffin and Skadburg (1969) suggested that above-chance hit rates of some studies could be due to scoring set. They claim that bias in the visual scoring method was due to the scoring set of looking for hits. Their study showed that scorers were indeed bias when they tried to count the number of hits in biorhythm curves.

This leads to the problem of defining a "hit," the point where the curve corsses the median. Should it be donfined to a period of 24 or 48 hours? Does the time of birth dias the data, etc? Previous studies assumed everyone was born at 12 noon, which is probably the only convention we can use. In this study we controlled for these problems by generating random birthdays and comparing them with biorhythms produced by actual birthdays.

Method

The subjects in this study were 852 personnel who were involved in industrial personal injury accidents, from 1973 to 1979, at an Air Force installation. The report of these accidents showed the accident date and employee number as well as other data on the accidents. The authors were able to secure the birthdays of all subjects from official records and then use them to calculate their biorhythms for the month of their accident. The computer plotted each of the three cycles for a one month period. This enabled us to visually inspect with a 24-hour size pattern overlay to determine if there were any "hits" during the day. This tedious task was accomplished by several people who were aware of the Chaffin and Skadburg article titled, "Effect of Scoring Set on Biorhythm Data."

Results

Table 1 shows a Chi-Square analysis of the number of hits for all three cycles for data generated by both actual and random birthdays. The random

Table 1

Chi-Square Analysis of Critical Day Accidents of all Cases with Biorhythms Computed from Actual and Random Birthdays (N = 852).

Cycle*	Expected Number Birthdays	Observed Actual Birthdays	X^2 Value Actual Birthdays	Observed Random Birthdays	X^2 Value Random Birthdays
P	74.01	70	.2172	59	3.0440
E	60.86	73	2.4216	55	.5642
I	51.64	57	.5563	59	1.0490
Total Chi-Square			3.1959		4.6572
Tabular Chi-Square, $p < .05$, df=2			5.9900		5.9900

*Physical (P), *Emotional (E), *Intellectual (I)

birthday group was produced by assigning random birthdays from the same general era to all subjects and recomputing their biorhythms for the month

of their accident. These plotted printouts were then inspected for the number of accidents which occurred on critical days and low points in the negative cycle. Table one shows that neither the actual or random birthday group reached statistical significance at the $p < .05$ level.

The fact that no significance was attained for the data produced by actual birthdays is significant in itself, but the randomized birthday data served as a double check for any systematic relationships in the data.

Table 2 shows a Chi-Square analysis of accident days that occurred at the lowest point of the negative half of the cycle. Here, both actual and random birthday plots reached statistical significance. This test, however, may not be creditable because of scoring problems. Theoretically, the low point should occur only one day of the cycle. But, since the plotted curve is very round at the low point, it is difficult to ascertain the exact nadir. Looking at the physical cycle only, the expected number for this measurement was based on 1/23(852) = 37.04. Interpretation of where the nadir occurs could be off more than a 24 hour period.

Table 2

Chi-Square Analysis of Low Point Day Accidents of all Cases with Biorhythms Computed from Actual and Random Birthdays (N = 852).

Cycle*	Expected Number Birthdays	Observed Actual Birthdays	X^2 Value Actual Birthdays	Observed Random Birthdays	X^2 Value Random Birthdays
P	37.04	44	1.3078	38	.0249
E	30.43	42	4.3991	47	9.0228
I	25.82	42	10.1391	42	10.1391
Total Chi-Square			15.8460		19.1888
Tabular Chi-Square, $p < .05$, df=2			5.99		5.99

*Physical (P), *Emotional (E), *Intellectual (I).

Perhaps a 1.5 or a 2 day range would be more appropriate. Then, the expected number would be more and would not reach statistical significance. Although the Chi-Square values are highly significant in Table 2, we must be hesitant to accept them as valid because of the scoring difficulties.

Table 3 shows the Chi-Square analysis of critical day accidents classified as "Unsafe Personal Factor: Inattentive, Absentminded." The Chi-Square analysis shows no statistically significant difference between the expected and observed number of hits. It was felt that this type of accident would be most likely to be related to biorhythm cycles.

Table 4 shows another likely to be related type of accident, that of "unsafe Practices: Unsafe Work Posture/Position." Here, 107 accidents

were identified and statistically analysized. Again, Chi-Square values did not reach significance at the p $<.05$ level.

Table 3

Observed and Expected Number of Critical-Day Accidents by Chi-Square Analysis of Unsafe Personal Factor: Inattentive, Absentminded (N = 128).

Cycle*	Exp. Number	Obs. Number	X^2 Value
P	11.13	14	.7400
E	9.14	13	1.6302
I	7.76	10	.6466
Total X^2			3.0168
Tabular X^2, p $<.05$			5.99

Table 4

Observed and Expected Number of Critical-Day Accidents by Chi-Square Analysis of Unsafe Practices: Unsafe Work Posture/ Position (N = 107)

Cycle*	Exp. Number	Obs. Number	X^2 Value
P	9.30	7	.5688
E	7.64	7	.1099
I	6.48	9	.9800
Total X^2			1.6587
Tabular X^2, p $<.05$			5.99

*Physical (P), *Emotional (E), *Intellectual (I).

Both of the previous accident catagories were also analyzed for low point accident relationships. The limited number of hits, however, both expected and observed, made an analysis less than reliable. The small numbers of 107 and 128 yeidled only expected numbers of hits at about three to five and Chi-Square analysis showed no significant relationship.

Discussion

In regard to accidents occurring on critical days in this study, the null hypothesis holds. The significance of the low point analysis is suspect, since the random birthdays also produced a similar significant relationship. Further evidence that there is no relationship between accidents and critical biorhythm days was demonstrated by the lack of a significant Chi-Square for the specific accident categories of unsafe personal factor and unsafe practices.

The authors agree with Palmer (1983); Shaffer, et al, (1978); and Wolcott, et al (1977) who suggest that several phenomena may be operating. Perhaps self-fulfilling prophecies, illusory correlations, sensitizations, concidences, selective reinforcement, and even the Hawthorne effect may be responsible for the popular belief in the theory.

Palmer (1983), a professor of physiology specializing in cyclic aspects of plants, animals, and humans, writes about science and biorhythm. He says, variability is the rule with human cycles; they also vary over time, with many modifying influences; there is no reason for them to start at birth and remain fixed over decades; and there is no reason for a critical point in a smooth transition from high to low. So, it seems that there is no "scientific" basis for biorhythms. Maybe if we could accurately measure an individuals recent cycles, we could check on the theory's validity.

References

Anderson, R. K. (1973). Man's timing mechanism. ASSE Journal, 18, 17-21.

Chaffin, R. & Skadburg,J. (1979). Effect of scoring set on biorhythm data. Journal of Applied Psychology, 64(2), 213-217.

Gittelson, B. (1975). Biorhythm, A Personal Science (2nd Ed.). NY: Arco Publishing Co., Inc.

Khalil, T. M., & Kurucz, C. N. (1977). The influence of biorhythm on accident occurrence and performance. Ergonomics, 20, 389-398.

Palmer, J. D. (1983). Fortunetelling revisited. Current, 249, 3-9.

Pittner, E. D., & Owens, P. (1975). Chance or destiny? A review of the biorhythm theory. Professional Safety, 20, 42-46.

Schwarz, G. R. (1976). A look at the matter of susceptibility to work errors as related to biorhythm. Professional Safety, 21, 34-39.

Shaffer, J. W.; Schmidt, C. W.; Zlotowitz, H. I.; & Fisher, R. S. (1978). Biorhythms and highway crashes. Archives of General Psychiatry, 35, 41-46.

Thommen, G. S. (1976). Is this Your Day? NY: Avon Books.

Thumann, P. E. (1977). Biorhythms and Industrial Safety. Atlanta, GA: The Fairmont Press.

Wolcott, J. H.; McMeekin, R. R.; Burgin, R. R.; & Yanowitch, R. E. (1977). Correlation of general aviation accidents with the biorhythm theory. Human Factors, 19, 283-294.

Management Information Systems: A Need For Human Factors

James A. Boyless
United States Air Force Academy
Colorado Springs, CO 80840

Abstract

The microcomputer revolution in management information systems now allows almost instant access to millions of bits of information to predict, trend, or even recall past activities. Additionally, the access to this information is now being accomplished by placing the computer on desktops everywhere. Often these computers are "user friendly" and little or no "computerese" is needed to operate them. These desktop computers have been human engineered from the software point of view but often the actual human engineering development and evaluation is lacking. This paper takes a general look at the lack of human engineering development and evaluation and suggests items that should be addressed in these two areas.

Introduction

Background

Revolutionary advancements in computer technology now allow individuals to readily access millions of bits of information. Additionally, these advancements now place the computer on desktops everywhere and there are very few places today where desktop computers are not used. The airline reservation you placed, the new car you insured, the automobile title you have, the automatic payroll deductions, and automatic transfer of funds are but a few examples of computer advancements. Furthermore, the integration of desktop computers in management information systems throughout the military sector allows instant access to information to predict, trend, or even recall past activities.

One management information system (MIS) in use at Headquarters, Tactical Air Command, is used to access and input information for all levels of management. The MIS is described as a "user friendly" system that is easy to operate with minimum of technical expertise. This is one type of system that has been "human engineered" from the software viewpoint. It has been designed for ease of use and little or no background requirements in "computerese" is needed. What could be added, however, is the actual human engineering development and evaluation between the hardware and individual.

Man-Machine Relationship

"Just what is he talking about?" may be the question in your mind right now. The answer is that there exists an intrinsic relationship between the hardware (computer terminal) and the individual (user). This relationship, called a man-machine system (MMS), is shown in figure 1. The MMS, studied in the field of Human Factors Engineering takes into account the premise that the best design enhances the effectiveness and efficiency of the intended goal of the system (McCormick & Sanders, 1982).

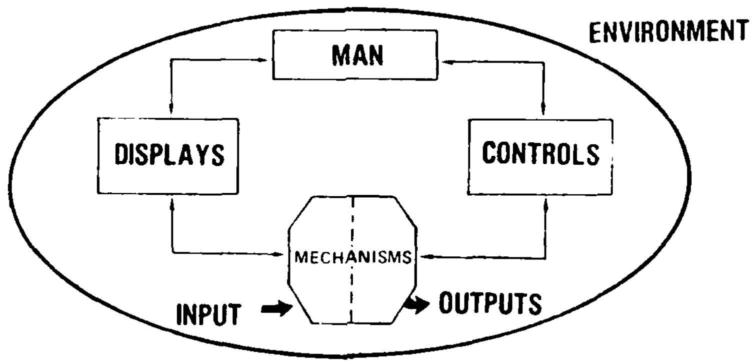


Figure 1. Schematic of Man-Machine System
Taylor, 1952

Man is shown as an integral part of the man-machine system (MMS). The MMS can be likened to many situations found everywhere. For example the MIS can be described as shown in figure 2.

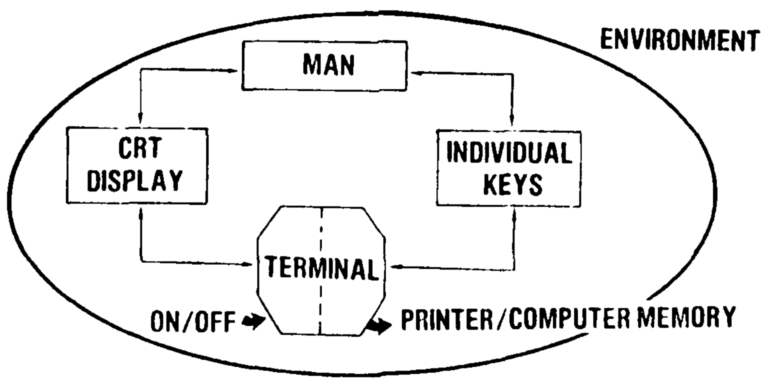


Figure 2. MIS Man-Machine Analogy

MMS Review

An explanation of the MMS can be perceived as follows. The display of the CRT serves as a stimulus for the man (operator) that triggers him/her to process the information presented (including decision making), which then prompts him/her to press the keys (control) to input the information. The entire process stated above is a closed-loop mechanical system because the effects of the machine (terminal) operation, as a result of the control action (pressing the individual keys), continually affects the information on the display (CRT commands and words) gathered in the associated work-station environment.

Although simplistic in explanation, the continuous interaction of man with his "work" tools has been an evolutionary task. As man happened against adversity, trial and error was the way things were designed and modified. But today, it is often costly to "retool" if design requirements are not met. Trial and error is no longer a valid method of design since one can imagine if the space shuttle was simply a trial and error project.

MMS Integration

But essentially trial and error is the method used in the design of workstations for management information systems. Very simply, the computer terminals are chosen after a careful system evaluation and ability to meet mission requirements and then placed on tables and desks meant to be used for other functions. The individual users are then expected to adjust to the situation at hand. If this is the "end" of system integration then the interrelationship of man and machine remains incomplete.

Completion of system integration depends on proper organization of the workstation. The workstation must be more than just a handy table or desk in the corner. It must take into account the ambient lighting

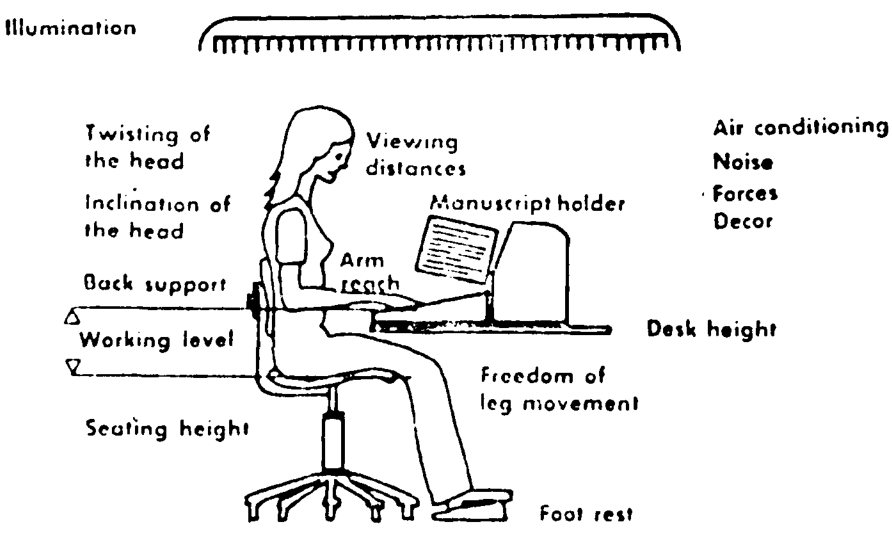


Figure 3. Main characteristics of a computer system workstation. (Chakir, 1980)

(including glare and contrast), associated furniture and equipment, and a number of other intrinsic factors in workstation design. Figure 3 shows the various details that must be studied in matching the design of the workstation to the task needs and physical characteristics of the individual.

Workstation Evaluation

Workstations have been evaluated in the areas of keyboard height, screen height, illumination, and glare. A number of design problems such as excessive keyboard heights and poor screen positioning were found. Poor screen positioning requires excessive head and neck inclination. In addition, operators were surveyed and results showed that there were a number of bothersome factors associated with viewing the screen, especially screen readability, reflected glare, screen brightness, and flicker (Stammerjohn, et al, 1981).

Lighting

Lighting is an area that needs to be closely evaluated in supporting computer systems such as a MIS. Proper illumination is essential so that both the computer terminal screen and resulting hard copy can be read without undue visual discomfort or fatigue (Stammerjohn, 1981). Unfortunately, specific levels of illumination are not available but guidelines and recommendations abound in the literature. The American National Standards Institute (ANSI, 1973) recommends minimum levels between 750 - 1600 lux[1] for general offices while other references recommend levels between 200 - 500 lux with supplementary task lighting. Lighting found in offices is generally sufficient but supplementary task lighting should be available. Glare, both direct and reflected, should be reduced or avoided. Direct glare, caused by light sources in the field of view, can be reduced through selection, number, and position of light sources. Reflected glare, caused by reflections of high brightness from polished or glossy surfaces and reflected toward the individual, can be reduced through low light levels, using indirect lights, and again position. Contrast levels are dealt with because the operator must be able to discriminate details of the printed matter being viewed (McCormick & Sanders, 1982).

Associated Furniture and Equipment

Associated furniture and equipment plays a big part in a properly designed workstation. Without proper regard for adequately designed furniture, the most expensive, elaborate, and "user friendly" computer system is purchased for naught. A computer workstation begins with the surface on which the computer terminal sits. This is a simple matter if a typing station is to be used since standard typing table heights are approximately 750mm, thus allowing the computer terminal home row to meet Mil-STD 1472C specifications for a sitting individual (see Table 1 for this and other recommended workstation dimensions).

---

[1] Lux is the luminous intensity/unit projected area of a surface in a given direction (Boyce, 1981).

TABLE 1. Recommended Workstation Dimensions

Name	Range	Source
Home Row Height	743-787mm	Mil-Std 1472C
Viewing Distance	438-502mm	"
Viewing Angle	10-20 Degrees	"
Arm Angle	80-120 Degrees	"
Wrist Angle	10 Degrees	Solomon & Burawa
Keyboard Height	720-750mm	Chakir
Knee Height Clearance	457-584mm	Van Cott & Kinkade
Chair, Adj Height	460-480mm	Stammerjohn

Surfaces should be directly adaptable to integrated computer terminals. If word and data processing is to be performed then a computer system table with adjustable height is recommended. The table should also be at least 750mm deep as a recommended workstation layout for either an integrated computer system or a two piece system.

There should be an adjustable video platform to allow the individual to place the video monitor at a comfortable viewing angle. An adjustable keyboard platform can be used with two piece computer systems allowing an appropriate home row height. A footrest can be adjusted for shorter individuals.

Conclusion

The inclusion of management information systems has allowed planners to have almost instant access to information for predicting, trending, and even recalling past activities. To accomplish the access of information it must be, of course, entered by individuals sitting at computer terminals. To insure good performance of these individuals, a human factors designed workstation is essential. Recommended dimensions of a typical workstation take into account: home row height, viewing angle and distance, arm angle, knee height clearance, and chair height among others. Not only are the dimensions of the workstation important, but also ambient lighting and other environmental factors. These factors should be evaluated in concert with any decision for management information system implementation. In all instances if the man-machine interface is carefully evaluated a more perfect union between machine and user can be achieved.

REFERENCES

(Due to space limitations, references have been omitted from this edition of the paper. References are available upon request from the author at the address given at the beginning of the paper.)

SYMPOSIUM LUNCHEON ADDRESS

PSYCHOLOGY AND BIG TIME INTERCOLLEGIATE ATHLETICS

Colonel John J. Clune
(United States Air Force Academy)

Responsible for the extensive Air Force Academy athletic and physical education program is one of the nation's top athletic administrators, Colonel John J. Clune, the sixth Athletic Director in Academy history. Colonel Clune assumed duties as athletic director on November 4th, 1975, after serving as assistant athletic director. During his tenure, he has guided a program that has met with excellent success. During the 1982-1983 school year, Academy varsity athletic teams won 64 percent of their contests. In the last five years, Falcon athletic teams under his direction have had outstanding records!

A 1954 graduate of the U.S. Naval Academy, Colonel Clune is no stranger to the world of athletics. While attending Navy, he earned All-American honors and set the Naval Academy single season and career basketball scoring records. The first record still stands and the career record lasted 25 years before being broken. He played on the Armed Forces Pan American team in 1955, was on the Air Force Olympic trails team in 1956 and coached several Air Force Base teams.

Colonel Clune received a master's degree in electrical engineering from the University of Southern California in 1964 and completed the Armed Forces Staff College in 1969 and the Industrial College of the Armed Forces in 1972. He is very active in collegiate national affairs, serving on the NCAA's Postseason Football Committee and serving as second vice president of the National Association of Collegiate Directors of Athletics.

His luncheon address at the Ninth Biennial Psychology in the DOD Symposium on 19 April 1984 is titled "Psychology and Big Time Intercollegiate Athletics."

PANEL SESSION

PANEL SESSION: COPING WITH MILITARY PERSONNEL TURNOVER: SOME APPLIED RESEARCH

SESSION CHAIR: H. Wallace Sinaiko (Smithsonian Institution)

PANELISTS: Kathleen Fernandes (Navy Personnel Research and Development Center)

Irwin G. Sarason (University of Washington)

David G. Bowers (University of Michigan)

Stuart A. Youngblood and Barry Baysinger (Texas A & M University)

PROCEEDINGS ENTRIES

"Coping with military personnel turnover: Some applied research" (Introduction - H. Wallace Sinaiko)

"Project RETAIN" (Kathleen Fernandes)

"Stress and stress coping: Reducing attrition among Marine recruits" (Irwin G. Sarason)

"Organizational factors in personnel attrition" (David G. Bowers)

"Behavioral and labor market correlates of personnel attrition and retention" (Stuart A. Youngblood and Barry Baysinger)

Coping with Military Personnel Turnover: Some Applied Research

H. Wallace Sinaiko
Smithsonian Institution

Introduction

High unscheduled personnel turnover--sometimes known as "attrition"--has been an unexpected consequence of the All-Volunteer Force. Although attrition rates are down from the unprecedentedly high levels of the mid- and late 1970's, losses continue to have serious consequences: They increase the costs of recruiting, perturb training schedules, and raise the costs of training; and, most importantly, losses can adversely affect readiness. Since about 1977 the DOD has supported a fair amount of research aimed at understanding attrition: why it happens and what can be done to prevent or reduce its occurrence. This panel includes four outstanding examples of applied research approaches to personnel attrition. Although each addresses Navy or Marine Corps populations, it is clear that research methods as well as findings can be extrapolated to other military settings.

One paper (Fernandes) deals with a particularly attrition-prone group of Navy people; the researchers eventually decided that a special training approach would be the best way to go, and they conducted field experiments to verify their hypotheses about ways to reduce turnover. The second effort (Sarason) takes the perspectives of organizational psychology and reports an investigation of the organizational determinants of involuntary discharges from the Navy. The third paper (Bowers) is about studies of attrition among Marine recruits and, in particular, the effects of stresses generated in boot camp; later phases of the research deal with the part played by drill instructors, with the personal stresses they experience, and with techniques that can be used to alleviate those stresses. The fourth study (Youngblood and Baysinger) analyzes turnover in a large cohort of Marine recruits over four years; both attitudinal and labor-market variables are among the significant predictors of attrition and reenlistment.

Project RETAIN

Kathleen Fernandes
Navy Personnel Research and Development Center

Project RETAIN is a two-part effort to reduce high attrition rates among general detail (GENDET) personnel. The underlying notion in RETAIN is to provide special training that will, in effect, get GENDETs "off to the right start" and thereby lower premature losses. There is no intention to keep people who are not suited to Navy life.

The first component of RETAIN is a three-day training program, known as REPORT, that is given to non-school-guaranteed recruits immediately upon their reporting to boot camp. REPORT focuses on inculcating skills and attitudes that will minimize the likelihood that bad attitudes and, ultimately, poor performance develop. REPORT is concerned with organizational assimilation, i.e., the means by which the various types of individuals available for military service can be most effectively integrated into a military organization.

The second component, FLOAT, provides three days of special training, either in a shipboard setting or in a classroom, at the point between standard apprentice training and first duty assignment. FLOAT emphasizes organizational commitment as well as the realities of life and work aboard ship. In FLOAT, the concern was to evaluate the effectiveness of establishing accurate initial expectations on later organization success.

Both REPORT and FLOAT were tested under field conditions. Each trial included control subjects, i.e., men, who were not put through either experimental training program. REPORT was evaluated in terms of its effects on performance in boot camp and on later attrition from the fleet; FLOAT was also assessed in terms of fleet attrition rates.

The results indicated that the REPORT participants had a lower training attrition rate and experienced fewer disciplinary problems than did control recruits. In addition, the REPORT participants who graduated from apprentice training maintained a lower fleet attrition rate than did controls after both groups had completed 32 months of service.

With respect to FLOAT, the shipboard version of the program had no effect on attrition. The classroom FLOAT program, however, had a substantial impact on reducing fleet losses. The attrition rate for the classroom participants was 13 percent lower than that for the control sailors after 29 months of service.

Stress and Stress Coping: Reducing Attrition Among Marine Recruits

Irwin G. Sarason
Department of Psychology, University of Washington

The first phase of this research dealt with the adjustment of Marine recruits. That work showed that from one-third to one-half of all attrition among trainees was due to behavioral causes. (There was no difference in performance scores of recruits who stayed and those who left, nor were age or prior education factors in attrition.) Researchers prepared special training materials intended to help recruits cope with the stresses of boot camp; the materials, in the form of TV tapes, encouraged realistic expectations that would increase recruits' confidence about meeting the challenges of basic training.

An unexpected finding was a wide variation in personnel loss rates among training platoons; e.g., attrition varied from zero to 28 percent. Further, there was an interactive effect: high-attrition platoons lost more non high school graduates, while education was not a factor in low-attrition groups. These results suggested to the researchers that drill instructors (DIs) were a key to understanding turnover among recruits.

Recent research has concentrated on drill instructors. These findings, among others, were reported: a) DIs who instill high personal trust tend to have low attrition rates; b) ratings by DIs at the end of boot camp are predictive of later performance; c) low-attrition training platoons tend to have DIs who see themselves as teachers, counselors, or parent surrogates; d) high-attrition platoons have DIs who believe in an authoritarian approach to leadership; e) a relatively high proportion of DIs are dropped from their school for psychological reasons, although the rates were inconsistent in 1980 and 1981; f) failed (i.e., dropped) DIs tend to be somewhat older than successful graduates; g) there are physiological correlates of DI performance; h) successful graduates of DI school have more realistic expectations of the rigors of training than those who fail; i) dropped DIs tend to have an external locus-of-control orientation vs. the opposite among successful DIs; and j) successful DIs tend to seek and use social support for their personal problems.

Continued study of DIs, some of whom are now 24 months out of their own training, has shown that they experience a great deal of pressure during the first year on the job. Many exhibit a steady increase in heart rate, an expressed preoccupation with speed, and impatience. These symptoms indicate, perhaps, a type of "burnout" not unlike that typical of teachers and nurses, among others. As a way of helping relieve some of the stresses experienced by DIs, the researchers have developed and are testing training films similar to those used with recruits; that is, DIs are taught the value of self-monitoring, how to control their anger, and how to deal with their worries about evaluations of their performance. New training modules are in preparation to enable DIs to handle special recruit-training problems and to observe recruits more accurately and provide feedback to them. Ultimately, all the training materials will be used routinely by the Marine Corps in a stand-alone mode, i.e., without any intervention by researchers.

Organizational Factors in Personnel Attrition

David G. Bowers
Institute for Social Research, University of Michigan

Project UPGRADE was an initiative of the Chief of Naval Operations to encourage commanding officers to discharge marginal performers. Subjects had to meet certain pay-grade and other criteria. UPGRADE was conducted during July-August 1981 and again in February-March 1982. The research examined organizational correlates of UPGRADE experience in terms of the distribution of Upgraders across unit types, reenlistment rates, unit readiness, and disciplinary infractions; there was also a series of case studies in which COs, division officers, work group supervisors and others were interviewed. The main findings were these: a) there was no relationship between unit size and the incidence of UPGRADE losses; b) Navy Human Resources Management Survey data obtained two to three years earlier did correlate with Upgrade; c) improvements in HRMS scores over time were predictive of lower Upgrade rates; d) reenlistment, UA, and desertion rates in a unit were related to Upgrade; e) there was no correlation between indices of unit readiness and Upgrade losses; f) drug usage was a factor only in the later (i.e., 1982) Upgrade exercise; and g) many Upgraders displayed anti-authority and anti-social behavior. The finding that organizational conditions were predictive of UPGRADE losses two or three years later, in spite of changes in COs and the usual rotation of personnel, led to the hypothesis that some people, in the wrong situation, become "vulnerable" to committing the misbehaviors that lead to UPGRADE discharge. There was also a hypothesis that "personnel velocity," or the sheer numbers of persons moving through a unit over time, is a condition that encourages misbehavior and leads to premature discharge.

Behavioral and Labor Market Correlates of Personnel Attrition and Retention

Stuart A. Youngblood
Barry Baysinger
College of Business Administration, Texas A&M University

During 1976-1981 the Center for Management and Organizational Research, University of South Carolina, conducted research on personnel attrition in the Marine Corps. There were three main components in the research: a) a longitudinal study of a cohort of 1500 men who entered the Marine Corps in 1976; b) additional studies of samples of men and women who entered the Marine Corps in 1977 and 1978; and c) the development and experimental evaluation of a realistic job preview intended to reduce attrition during recruit training. This report deals only with the first component.

The longitudinal study, as indicated, tracked an entering cohort of about 1500 marines for four years as a way of relating changes in perceptions, expectations, values, and intentions to attrition and reenlistment. Data on the cohort were collected at several stages: pre-recruit training, post-recruit training, and after assignment to first duty station. The data collected included: individual measures (e.g., personal attributes, role attractions of Marine Corps and of civilian jobs), organizational measures (e.g., leadership, job characteristics), and criterion measures (e.g., intentions, performance, and attrition). The research sought to distinguish three groups from one another: a) early leavers (i.e., attriters during recruit training), b) later leavers (i.e., attriters after recruit training), and c) stayers (i.e., men who completed their initial obligated service). There was also a focus on changes in later leavers and stayers over time, and on men who reenlisted at the end of their tours. Among the findings of this part of the research were: a) men who attrited were less well educated, and they performed less well on entry tests of mental ability; b) all classes of subjects consistently exhibited favorable attitudes toward the Marine Corps at the end of recruit training, and all showed sharp declines thereafter; c) men who entered with four-year obligations had "initially higher completion and reenlistment intentions, higher attraction to the military, higher internal motivation and growth need strength, higher expected satisfaction, and more favorable job and work group perceptions." Good predictors of whether men would stay in the Marine Corps or leave before their tours were measures, collected prior to the start of recruit training, of subjects' stated intentions to complete their enlistments and to reenlist.

Comparisons of significant predictors among attitudinal and labor market variables for early turnover were made, as were comparisons of influences on later turnover and reenlistment decisions. (Findings from this phase of the research will be reported to the panel.)

PANEL SESSION

Emerging Issues in Aerospace

Team Training

Participants: Larry Liberty
President, Liberty and Associates
Organizational and Management Consultants

Robert Helmreich
Professor of Social Psychology
University of Texas

Michael Sparks
Professor of Business
California State University, Sacramento

Trieve Tanner, Ph.D.
Chief, Space Human Factors Office
NASA-Ames Research Center

Background

As we enter the middle of the 1980's we continue to see an expanding amount of interest in the use of "team training" and the related technologies for expanding productivity. As far back as 1927 when the famous "Hawthorne Studies" began, there began to be an interest in discovering how to have people produce more and do it in a way that did not endanger the organization. (Dickson and Roethlisberger, 1939)

In the late 1930's Kurt Lewin began to expand the reach of the psychological sciences to the study of groups and their behavior. This was a profound evolutionary step in organizational interest and awareness of the implications of groups, their behaviors, social dynamics, et al. Lewin successfully demonstrated that groups are "living systems", and thus created the first contemporary paradigm for the creation and study of group behavior. (Lewin, 1945, 1951)

Since that time the creation of "social psychology," "organizational behavior," and the "behavioral sciences" have become common words in the lexicon of managers, academics, and all of the fields and profession closely associated with organizations. In fact these words are now part of the normal organizational jargon for any "up-and-coming" junior executive.

As the last fifteen years have passed a number of new developments have influenced the current status of team related training and development:

1) The creation of "organizational development" as a discipline

and profession

2) The growing acceptance, tolerance and useage of persons who call themselves "consultants"

3) The expanding importance of "interpersonal" dynamics and issues in space/aerospace

4) The growth of high technology in all of its forms and uses in aerospaces milieus

5) A growing awareness and acceptance of the importance of the individual worker, and the work relationships upon quality/quantity of results

The importance of these developments can only be seen in the context of how important we now feel "team training" is. In nearly every organization of any magnitude we now have trainers who either do training related to teams and groups, or these trainers screen, recommend, or hire professional consultants who espouse various specialties related to teams and their effectiveness. Even our linguistics have succumbed to this emerging profession. O.D. consultants call themselves "practioners," a term we have always used with care to reflect members of the medical field.

Current Situation

In the last few years the airlines industry has begun to study and work more vigorously in the area of team performance productivity and effectiveness. In June of 1979 NASA, in conjunction with the airlines industry, held a workshop oriented around beginning to thoroughly look at some of the critical issues in this area. Its title, "Resource Management of the Flight Deck" indicates the emphasis and concern about having airline crews begin to work more effectively together.

Several airlines have voluntarily developed and implemented trainings in this area. United Airlines has a special four day seminar for all pilots which is designed to expand both self-awareness and skill in working together in the cockpit. And yet there is growing concern that this may not be enough to really impact the situation.

Numerous accidents and "incidents" have begun to be more closely tied to the impact of crew "team relationships" upon the safety and behavior of those persons responsible for flying the aircraft. The literature about airline safety/accidents is full of examples where there have been "breakdowns" amongst the cockpit crew. In some cases there is ample evidence that this failure to work well together cost lives as well as aircraft.

In this context there is growing interest in determining just how important, effective, and useful our current approaches to team training and development are. As more and more attention is given to this subject we are beginning to see several new and critical issues emerge for consideration.

Some Emerging Issues

As institutions and organizations have continued to look for answers to problems associated with individual and group perfomance and behavior, and as more professionals, academics, and scientists have begun to attend to these areas, new issues and concerns have begun to surfact. Some of the most current one include the following:

1. Rethinking the Abstractions of "Team"

What is the meaning, definition, and fundamental nature of "teams"? What are the underlying concepts and abstractions from which we give meaning to this word? It is becoming clearer that what one person calls "team" building may be totally different from the next. Is there such a quality as "teamness" or "teamship" and how are such characteristics or qualities defined so as to be useful to all parties who have interest?

2. Rethinking the Abstractions Underlying Team Training

Given the question of what "team" or "teamness" may be, it creates a new series of questions and issues related to the nature of current "team training" technologies. Are we clear about the content, context, and process needed to foster effective learning and development in this area? At which levels do we develop technologies appropriate to ensure such learning-individual, relationship, group? Are the learning paradigms and abstractions that we currently have in use in "team building" and associated approaches actually appropriate once we decode the definitions of "team" et al.?

3. Identifying Real and Potential Results of "Teams"

A new and still unanswered question relates to whether every group that works together or associates closely should be worked with to be a "team." It is possible that some groups may not need such development and if so where and how do we identify these scenarios? Are there "negative" results that may accompany team training that should be considered in the broader context of organizational and individual performance? A variant issue is to determine what a team can produce in results vis a vis a task force, individual, or well-managed group.

4. The Implications of "Boundary" Definitions on Teams

An emerging question when this "teamness" quality does show up is "where is the "inside" and where is the "outside" of any particular team?" How

does the creation and location of this boundary effect effectiveness and productivity? How do these boundaries come into existence and what is their fundamental nature and composition?

These are some of many important and interesting issues beginning to surface as we engage the question of how to produce breakthroughs in having people work together productively.

Additionally there are a number of related issues in the domain of selection. These tie together with these issues of team training in that we may find that certain personality types work more effectively together. Selection may become as critical to the ultimate effectiveness of a group's results as the manner in which the group is managed, trained and developed, or organized.

Selected Bibliography

Argyris, Chris. Integrating the Individual and the Organization New York: Wiley, 1964

Bennis, Warren G. Changing Organizations New York: McGraw-Hill, 1966

Bennis, Warren G., Benne, Kenneth D., and Chin, Robert, eds. Planning of Change: Readings in the Applied Behavioral Sciences. New York: Holt, Rinehart & Winston, 1961

Lewin, Kurt. A Dynamic Theory of Personality. New York: McGraw-Hill, 1945

Lewin, Kurt. Field Theory in Social Science. Selected Theoretical Papers New York: Harper & Row, 1951

Liberty, Larry L. Teams in Space: A Proposal for an in-depth Survey of "Teams," "Team Building," and "Team Performance;" A Proposal to Ames Research Center, 1983

Roethlisberger, F.J.; and Dickson, William J., Management and the Worker, Cambridge, Massachusetts, Harvard University Press, 1939

Sehon, Edgar H. Organizational Psychology, Englewood Cliffs, New Jersey; Prentice-Hall, 1965

PANEL SESSION

PANEL SESSION: INDIVIDUALIZED INSTRUCTION

SESSION CHAIR: Barbara L. McCombs (Denver Research Institute)

PANELISTS: Barbara L. McCombs, Dale A. Steffen, James P. Kottenstette, J. Jeffrey Richardson, Louis F. Cicchinelli, and Anita S. West (Denver Research Institute)

PROCEEDINGS ENTRIES

"Individualized instruction" (Program summary)

"Individualizing with computer-assisted instruction (CAI): Enhancing student motivation through computer-controlled audio"(Barbara L. McCombs)

"Training and performance support system: A new dimension in individualized instruction" (Dale A. Steffen and James P. Kottenstette)

"Maintenance troubleshooting simulation: Adapting to the trainee" (J. Jeffrey Richardson)

"Evaluation of the electronic equipment maintenance trainer (EEMT)" (Louis F. Cicchinelli)

"Factors influencing the selection of individualized instructional approaches" (Anita S. West)

Individualized Instruction

Denver Research Institute

University of Denver

This panel focuses on progress made and research issues remaining in a number of areas generally related to individualized instruction in military training. The following abstracts describe the five topic areas and list the Denver Research Institute (DRI) personnel presenting each topic in the panel.

Maintenance Troubleshooting Simulation: Adapting to the Trainee

J. Jeffrey Richardson, DRI/Social Systems Research and Evaluation Division

Maintenance troubleshooting tasks certainly involve more than equipment manipulations. Behind each manipulation is a series of mental decisions which establish its purpose. Therefore, a maintenance troubleshooting simulation must model not only equipment manipulations, but also the cognitive component of the task. Once such a model has been developed and used to generate a series of troubleshooting problems, it is possible to adapt the way the simulations interact with the trainee in ways that focus attention on unmastered components of the task, thus increasing training effectiveness and efficiency. For example, simulations often offer a menu of alternatives at each decision point. When a *trainee demonstrates competence in selecting the correct option from the menu*, the task may be made slightly more difficult by not displaying the menu. Now the trainee is expected to recall the response set from memory and then select the proper alternative. As a second example, when a complex procedure is simulated, task-centered simulations require the trainee to explicitly state plans and goals leading to equipment actions before performing them. Once individual equipment actions are correctly performed, punctuated by explicit verbal statements of intent, the trainee may be allowed to "chain" together commonly co-occurring equipment actions into smooth sequences without intervening verbal mediation. This presentation will describe research on maintenance simulation funded by the Air Force Human Resources Laboratory which has produced a maintenance simulation system with the capability of adapting to trainee growth in competence in a way that is sensitivve to differential rates of growth in competence in each subcomponent of the task. This adaptability is achieved through the use of an on-line task analysis which serves as an overlay model of the trainee.

The Training and Performance Support System: A New Dimension in Individualized Instruction

Dale A. Steffen, DRI/Electronics Division
James P. Kottenstette, DRI/Social Systems Research and Evaluation Division

A project was initiated in 1981 by the Denver Research Institute to develop a prototype Training and Performance Support System (TPSS). The TPSS is a computer-based training and job-aid system scheduled for delivery in mid

February 1984 to the Electronic Systems Division of the Air Force Systems Command at Hanscom AFB, MA. In addition to hardware and software components of the TPSS, instructional materials have been developed to train and aid Air Force computer specialists responsible for managing the acquisition of computer software for major defense systems. The system was designed to provide a capability for increasing productivity of professional personnel who assume new and different job responsibilities. Such design includes a technique which links job preparation (training) with job performance (productivity) by; (1) permitting training that is based on job tasks, and (2) providing routine access to an instructional resources data base during on-the-job activities. The data base provides task-oriented job aids and will be delivered via existing networks in the work environment. This technique allows a capability for workers to integrate on-the-job performance aids with routine data in the work environment to produce products in a more effective and efficient manner. The combination of use on demand and the integration of instruction and information provides a new dimension in individualized instruction. The status of instructional materials, software, and hardware components of TPSS will be presented, along with a discussion of implications of this approach for individualized on-the-job training.

Evaluation of the Electronic Equipment Maintenance Trainer (EEMT): Individualized Troubleshooting Training

Louis F. Cicchinelli, DRI/Social Systems Research and Evaluation Division

The EEMT system is currently being used to train Class "A" School students to maintain electronics equipment. These prototype devices have been installed in the Electronic Technician (ET) School at the Service School Command at Great Lakes, Illinois, and are used to simulate radar and electronic test equipment. The trainer is designed to provide students, on an individual basis, with experience in troubleshooting radar system malfunctions using videodisk equipment displays. This presentation will include a brief description of the EEMT trainer and the evaluation methodology, together with a discussion of the evaluation findings in five areas: training effectiveness, subsequent student performance, cost, suitability and user acceptance. Implications of the system for providing individualized troubleshooting training will also be discussed.

Individualizing with Computer-Assisted Instruction (CAI): Enhancing Student Motivation Through Computer-Controlled Audio

Barbara L. McCombs, DRI/Social Systems Research and Evaluation Division

A computer-controlled audio capability for enhancing the effectiveness of motivational skills training for military trainees was developed for the Army Research Institute as part of their Basic Skills/Learning Strategies research program. This capability interfaces with an Apple IIe microcomputer system and provides for the personalization of computer-assisted introductory and practice segments for seven printed, self-instructional motivational skills training models. The modules promote the development of self-management, personal responsibility, and positive self-control skills which underlie self-motivation. The implications of this research for providing a cost-effective technology for personalizing CAI

training and for reducing the dependence of the motivational skills training on instructor and group process facilitation are discussed.

Factors Influencing the Selection of Individualized Instructional Approaches

Anita S. West, DRI/Social Systems Research and Evaluation Division

Recent research for the Air Force Human Resources Laboratory has focused on the selection and implementation of self-paced resident training methodologies, including factors that influence the selection of computer-based technologies and factors impacting the successful implementation of nonconventional instructional formats. The results of these research programs are consolidated and presented in the form of a model for decision makers faced with the decision to individualize their resident training programs.

AD-P003 314

Individualizing with Computer-Assisted Instruction (CAI): Enhancing Student Motivation Through Computer-Controlled Audio

Barbara L. McCombs, Ph.D.
Senior Research Scientist
Denver Research Institute

Abstract

A computer-controlled audio capability for enhancing the effectiveness of motivational skills training for military trainees was developed for the Army Research Institute as part of its Basic Skills/Learning Strategies research program. This capability interfaces with an Apple IIe microcomputer system and provides for the personalization of computer-assisted introductory and practice segments for seven printed, self-instructional motivational skills training modules. The modules promote the development of self-management, personal responsibility, and positive self-control skills that underlie self-motivation. The implications of this research for providing a cost-effective technology for personalizing CAI training and for reducing the dependence of the motivational skills training on instructor and group process facilitation are discussed.

Introduction

Many students entering military technical training lack not only prerequisite basic reading skills and cognitive learning strategies, but also demonstrate skill deficiencies of an attitudinal or motivational nature. Although some attention has been given to programs that remediate reading skills, cognitive learning strategies, and study skills, little attention has been given to programs that can remediate the strategies and skills related to trainee motivation. The development and implementation of such a program promises to improve the military trainee's ability to positively adjust to the requirements of military technical training through the acquisition of a variety of self-management, personal responsibility, and self-control strategies which can increase trainee motivation. The program also has the potential of reducing the high costs associated with eliminating motivationally deficient trainees after they have completed sizeable portions of technical training.

A program for accomplishing the preceding goals in military technical training, entitled the Motivational Skills Training Package, has been developed and recently evaluated in a contract for the Defense Advanced Research Projects Agency (DARPA). The program includes seven self-instructional, printed modules that have been implemented in an instructor-led, small-group format which provides trainees with the opportunity to practice new strategies and skills, share experiences, and develop feelings of rapport with their instructors and fellow trainees. The evaluation of this program of self-instructional materials augmented by instructor support and group experiences in the Air Force's Precision Measuring Equipment (PME) course indicated that (a) trainees liked the program and found it helpful in both their coursework and personal lives and (b)

trainees participating in the program had significantly higher block test scores and lower block test failure rates than control group trainees (McCombs & Dobrovolny, 1982).

Although the evaluation findings with the Motivational Skills Training Package clearly pointed to its success, several research questions remained. One set of questions concerns the issue of the format of this training package and the use of instructors and the group process to facilitate trainee acquisition and maintenance of strategies and skills included in the package. For example, could the cost effectiveness of the program be enhanced by reducing instructor-trainee requirements and/or group interaction requirements through the use of computer-assisted instruction (CAI) for selected portions of the training? An investigation of this question is currently being funded by the Army Research Institute (ARI). To date, CAI enhancements have been developed to augment the printed package and an evaluation study is now underway with military trainees at Ft. Sill's Electronics Communication School. The purpose of this presentation is to describe the design of CAI materials and the implications of the resulting technology for personalizing CAI training. As an introduction to these sections, the next section will briefly describe the content and strategies being taught in the Motivational Skills Training Package.

Motivational Skills Training Package

The content and structure of this motivational package was defined as a result of an in-depth experimental analysis of specific conative, affective, and cognitive skill deficiencies of Air Force trainees in four technical training courses at Lowry Air Force Base (McCombs & Dobrovolny, 1980). This analysis indicated that the primary deficiencies (i.e., characteristics that differentiated effective from ineffective learners) in the conative domain were that poorer students consistently had low motivation to learn, had few military or personal goals, could be classified as being low in maturity, with little self-discipline or the ability to take responsibility for their own learning. In the affective domain, poorer students were generally those with high levels of anxiety toward learning and taking tests, and who lacked effective skills for coping with the demands of technical training. In the cognitive domain, the poorer students were generally those with poor reasoning and comprehension skills, and/or those who lacked effective decision making and problem solving skills in technical or personal areas.

Based on this analysis which integrated relevant literature, student performance data, student and instructor interview results, and individual difference data, seven skill-training modules were defined (McCombs, 1982a, 1982b). The Introduction Module introduces students to the concept of personal responsibility and positive self-control, presents rudimentary techniques for controlling negative attitudes (e.g., use of positive self-talk and imagination), and explains the purpose of the skills training package. The Values Clarification or Self-Knowledge Module explains the role of values and beliefs in helping us define ourselves and what's important to us, stresses each person's responsibility in defining his or her own value system, and helps students explore their values and beliefs in a number of areas. The Career Development Module builds on students' newly acquired self-knowledge and helps them acquire the necessary decision-making skills to explore their career interests and make some career goals and plans. The Goal-Setting Module formalizes the previously learned goal-setting process by first describing the purpose of goals as directing and motivating human behavior, describing a general model for systematically thinking about and setting personal goals, and helping students work through exercises for setting specific long-term and short-term goals. The Effective Communication Module describes techniques and strategies for

effectively communicating feelings, wants, and needs and for dealing with stressful interpersonal situations that may impede goal attainment. The Stress Management Module describes the role of perceptions, negative self-talk, and mistaken beliefs in producing stress and presents a number of generalizeable strategies for managing stress. Finally, the Problem Solving Module provides a summary of the skill-training package by pointing out that students have been using a problem solving approach throughout this training and by providing a general model for systematically working through and solving personal and technical training problems.

Design of CAI Enhancements

The objectives of the design of CAI enhancements for the Motivational Skills Training Package were to identify those components of the training that were enhanced by instructor and/or group experiences (McCombs & Lockhart, 1983). An analysis of the instructor's functions in facilitating student acquisition of strategies and skills taught in the package indicated that the instructor was instrumental in establishing a good relationship or personal rapport with the student, in serving as a model of personal responsibility and positive self-control, in helping the student understand what is expected of him or her in the training program, in introducing important concepts in each module in order to provide an advance organizer or meaningful structure for acquiring new concepts and skills, and in reinforcing the value of the skill training for positive self-development by explaining its application and benefits in military experiences. Group process components identified as facilitative included helping students identify with peers and open up to sharing personal feelings and experiences, providing opportunities for shared decision making and friendships, helping students reinforce mastery of new skills by group rehearsal and feedback, and assisting students in behavioral assignments and contracts that promote skill maintenance after the training is over.

Based on this analysis of facilitative instructor and group process functions, elements that could be simulated by specific CAI interactions were identified. In the area of instructor functions, three primary roles that could be provided by the computer were defined: modeler, facilitator, and motivator. To provide these roles, a character named "PC" was created to serve as an instructor/guide and to interactively perform each role by demonstrating the use of new strategies and skills, providing introductory concepts in a meainingful context, and coaching students in the application of new concepts and skills via personalized feedback and encouragement. In the area of group process functions, a set of military trainee characters was created to represent specific personal responsibility/self-control problems related to each module's content area. That is, a male or female character and accompanying problem scenario was defined to exemplify typical student problems with personal responsibility in general, with knowing who they are and what's important to them, with knowing their career interests and goals, with knowing how to set goals, with knowing how to manage stress, with knowing how to communicate effectively, and with knowing how to solve problems. These characters were designed to "grow" as a result of their skill training from their initial inability to solve particular problems to competent problem solvers and self-managers. This transition was designed to occur between PC's guided CAI introductions and CAI practice sessions for each module. CAI segments were therefore designed to incorporate these instructor and group process elements by providing introductions to each printed module as well as practice sessions following student reading of each module.

To achieve the high degree of personalization required to make PC and the seven student characters highly realistic and easily identifiable necessitated the use of a rich

training medium that could provide both interactive visual and auditory capabilities. Although videodisc technology is available to provide these capabilities, it was beyond the scope of the contracted effort to interface a videodisc to the project-purchased Apple IIe systems, nor was this level of sophistication deemed necessary to provide the type of personalization desired. A search was thus begun to locate a less costly audio cassette player that could be interfaced to the Apple IIe microcomputer. The search revealed no such product yet on the market and work then began on the design and development of a computer-controlled audio capability.

The resulting capability consists of a specially designed interface card which plugs into the Apple IIe game I/O port. The interface receives pulses from a standard slide-sync audio cassette player. These pulses are used to trigger screen changes and, in turn, to allow the CAI software (in this case, the Apple SuperPILOT Authoring System) to control the on/off function of the audio player. This capability allows for computer control of a linear sequence of audio messages that coincide with particular CAI frame sequences, as well as provides for the personalization of skill training introductions and practices, at about one-eighth the cost of videodisc technology. In this particular application, audio is integrated with screen information in the following ways: (1) CAI screen reinforcement of audio information, (2) audio instruction preceding a CAI segment, (3) audio feedback following a CAI segment, and (4) audio as an integral part of a CAI segment.

Implications of the Computer-Controlled Audio Technology

Preliminary evaluation results at Ft. Sill from the implementation of the computer-controlled audio enhancements to CAI segments which accompany the Motivational Skills Training Package indicate that this technology is being positively received by both students and instructors and is reducing instructor and group interaction requirements. Although final evaluation results comparing the effectiveness of the CAI enhancements with actual instructor-led introductions and group practice sessions are still forthcoming, it is possible to speculate about several important implications of this technology for providing a cost-effective approach to the personalization of CAI training.

First, the benefits of the technology for this type of skill training are that it can increase personalization, allow simulation of instructor and group process functions, introduce novelty to enhance student motivation, encourage students to maintain attention to relevant screen information, boost skill levels regardless of reading ability, reduce reading demands as well as potentially improve reading skills directly, enable training by example (e.g., in the communication skills area), provide consistency in training, and save production time compared to other media. Many of these benefits are applicable to skill training domains outside the motivational area, thus suggesting the application of this technology in the training of cognitive and metacognitive learning strategies (e.g., reasoning and memory skills, self-assessment and self-monitoring skills) and in the training of specific cognitive and procedural skills required in technical training.

Second, in order to fully capitalize on the potential benefits of the technology in other training areas, a programmatic research effort is suggested. The research questions to be addressed in such a program include first assessing the instructional benefits of audio enhancements to CAI in the preceding areas (i.e., does audio improve training effectiveness?). Once the benefits are established for each area, the degree of complexity and/or individualization required to maximize training effectiveness can be

addressed (i.e., what degree of complexity/individualization is most beneficial in each training area?).

After answers to the preceding research questions have been obtained, a third area to be explored is the technological advancements that can improve the computer-controlled audio capabilities currently developed. Areas where improvements are needed include the speed of use and ease of use for the user, the complexity of the training strategies and/or individualization that can be supported by the technology, and the ease of audio production possible with more sophisticated addressable and random access capabilities. For example, it may be desirable for some skill training applications to have a forward or backward branching capability, to have increased timing accuracy and precision for screen/audio information sequences, and to increase the audio quality by eliminating pulse sounds. The important thing to keep in mind, however, is that the instructional need for technological advancements be established prior to investing in the necessary development work.

In summary, there appear to be many skill training areas in which a computer-controlled audio capability can enhance both the personalization and individualization of training, thereby also enhancing student motivation. This technology is seen as a more cost effective approach than videodisc technology, particularly because a wide range of training applications do not require the use of both the CAI and video media in conjunction with audio. Further exploration of the technology thus promises to expand the range of options available to us for individualizing with CAI using currently available microcomputers.

References

McCombs, B.L. Transitioning learning strategies research into practice: Focus on the student in technical training. Journal of Instructional Development, 1982, 5(2), 10-17. (a)

McCombs, B. L. Learner satisfaction and motivation: Capitalizing on strategies for positive self-control. Performance and Instruction, 1982, 21(4), 3-6. (b)

McCombs, B. L., & Dobrovolny, J. L. Student motivational skill training package: Evaluation for Air Force technical training (AFHRL-TP-82-31). Lowry AFB, CO: Air Force Human Resources Laboratory, Technical Training Division, December 1982.

McCombs, B. L., & Dobrovolny, J. L. Rationale for the design of specific types of student skill training in the conative, affective, and cognitive skill domains (AFHRL Interim Technical Report). Lowry AFB, CO: Air Force Human Resources Laboratory, Technical Training Division, June 1980.

McCombs, B. L., & Lockhart, K. A. Motivational skills training: Helping students adapt by taking personal responsibility and positive self-control. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada, April 1983.

Training and Performance Support System: A New Dimension in Individualized Instruction

Dale A. Steffen and James P. Kottenstette
Denver Research Institute

Abstract

A project was initiated in 1981 by the Denver Research Institute to develop a prototype Training and Performance Support System (TPSS). The TPSS is a computer-based training and job aid system scheduled for delivery in mid-February 1984 to the Electronic Systems Division of the Air Force Systems Command at Hanscom AFB, MA. In addition to hardware and software components instructional materials have been developed to aid Air Force computer specialists responsible for managing the acquisition of computer software for major defense systems. The system was designed to provide a capability for increasing the initial productivity of professional personnel who must assume new job responsibilities. The design concept links job preparation with job performance by: 1) providing instruction for and an explanation of major job tasks, and 2) providing routine access to task-related information for use during job performance activities. This technique provides workers with "performance aids" in the work environment that are an extension of the instructional content structure. The combination of use on demand and the integration of instruction and information provides a new dimension in individualized instruction through learner control.

BACKGROUND

The TPSS application concerns the preparation and support of adult learners for new roles as managers in the acquisition of major defense systems. Broadly stated, the role of Air Force acquisition managers is to control the technical risks associated with the development and production of new equipment. Failure to control these risks typically results in higher system costs and deferred system availability.

The novice manager's role is especially difficult because: the systems are enormously complex, requiring extensive teamwork among functional area managers; management practices are obscured by a plethora of policies, regulations, standards and procedures; and finally, novice managers have limited opportunities for formal training-- experience is the teacher and the Program Office is the classroom.

From an organizational perspective, the need to provide some type of support for less-experienced technical managers is acute: for example, 25% of the military work force is reassigned each year and replacement personnel

typically have little acquisition management experience. Providing cost-effective support, however, is not a simple thing. The traditional approach, suggested in Figure 1, is to establish a training organization, develop knowledge and skills-oriented training components, and graduate novices ready for Program Office assignments. The underlying assumption is that with time, experience, and additional learning, the novice will become expert.

The traditional approach to training works well when many people are expected to perform similar tasks. It works less well for small groups of people that must perform many different tasks because the training establishment necessary will overwhelm the host organization and mission. Thus, in the past, the default position has been to staff each Program Office, "let experience be the teacher," and simply not seek "expert" behaviors from novices.

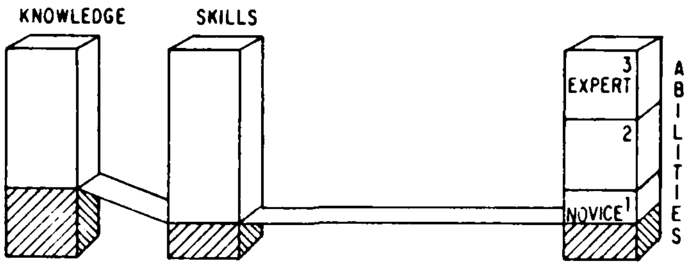


FIGURE 1 THE TRADITIONAL VIEW: TRAINING/JOB PERFORMANCE

The Training and Performance Support System (TPSS) was invented to meet an urgent organizational requirement for "expert" performance by novice and less-experienced personnel. TPSS addresses this inherent contradiction through a three-tiered structure for instruction and information named, respectively, the Education, Job Entry, and Performance Support Levels. The Education Level of TPSS provides for a classroom-based orientation to acquisition management. Learning at this level ignores individual differences: it serves to qualify persons for individualized instruction. The Job Entry Level provides computer-based instruction so that a unique curriculum can be tailored for each person, based on "events" or "milestones" scheduled to occur in the acquisition life cycle of his/her program assignment. The idea is to provide novice personnel with in-depth instructional resources, initially utilized on a scheduled basis, and on a demand basis as work progresses. The Performance Support Level provides computer-based access to information known to be relevant in performance of job tasks that surround each event. It is the imposition of content structures on this information that makes it possible to: 1) develop individualized instruction, and 2) organize "performance aiding" information that fosters "expert" behavior.

Figure 2 indicates how expert behavior is supported by the TPSS concept: knowledge and skill-oriented learning components implemented at the Job Entry Level are shown to be integrated with the information structures from the Performance Support Level. Since the learning components encourage problem-solving and information seeking behavior, the essential elements

believed necessary to foster expert job performance are present, albeit tne expertise is limited.

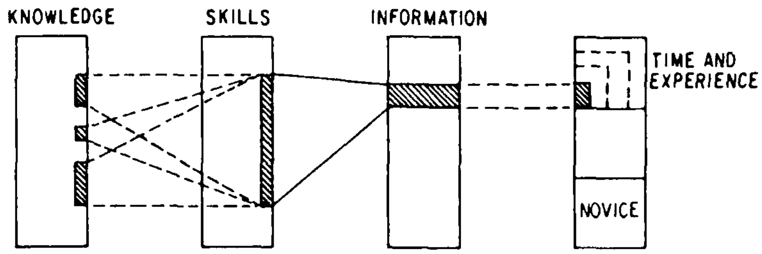


FIGURE 2. TPSS VIEW OF JOB SUPPORT

A NEW DIMENSION FOR INDIVIDUALIZED INSTRUCTION

If individualized instruction is taken to mean that instructional transactions are somehow modified in light of learner characteristics, then TPSS establishes a new dimension in individualization. The new dimension is achieved through learner control over the depth of instruction during each learning episode, and more generally, through learner control over the knowledge -- information continuum associated with each Job Entry Lesson.

In the latter instance, control is given to system users through their familiarity with the content structures imposed on the instruction and information prepared for each event in the acquisition process. Since these structures are derived from engineering management functions, and are consistently applied to the instruction and information, they are readily mastered by learners. Figure 3 shows these structures for a lesson and for its corresponding Performance Support information base.

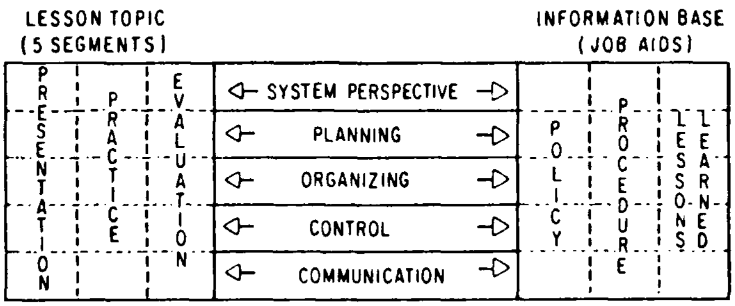


FIGURE 3. INSTRUCTION AND INFORMATION STRUCTURES

Notice that each lesson is composed of five segments or modules, and each segment treats a particular aspect of the event. For example, if the event is an engineering design review, the SYSTEM PERSPECTIVE SEGMENT "locates" the review within the acquisition life cycle and introduces the management principles that apply; the PLANNING SEGMENT considers scheduling and coordination considerations; the ORGANIZING SEGMENT addresses how the government manager prepares for participation in the review; the CONTROL SEGMENT deals with the dynamics of the review itself; and the COMMUNICATION

SEGMENT develops interactions and documentation of results. Taken together, the lesson segments provide the novice with a simulation of the event as well as a framework for treating many different knowledge requirements.

The figure also shows the information base structure for such events. In addition to the segment classifiers, the information is also classified by type: POLICY, PROCEDURES, and LESSONS LEARNED. This two-way classification of information facilitates its use as a job aid, independent of instructional considerations.

Learner control over the depth of instruction is achieved through the instructional model itself. As illustrated in Figure 4, each lesson segment has one or more MAIN IDEAS that must be mastered, and each idea is developed through several GENERALITIES. The Generalities represent one level of content treatment selected by the learner. ELABORATION or expansion on each generality represents deeper development of the idea; and lastly, the learner may chose to interrogate the information base from the lesson, and this exhausts learner control possibilities.

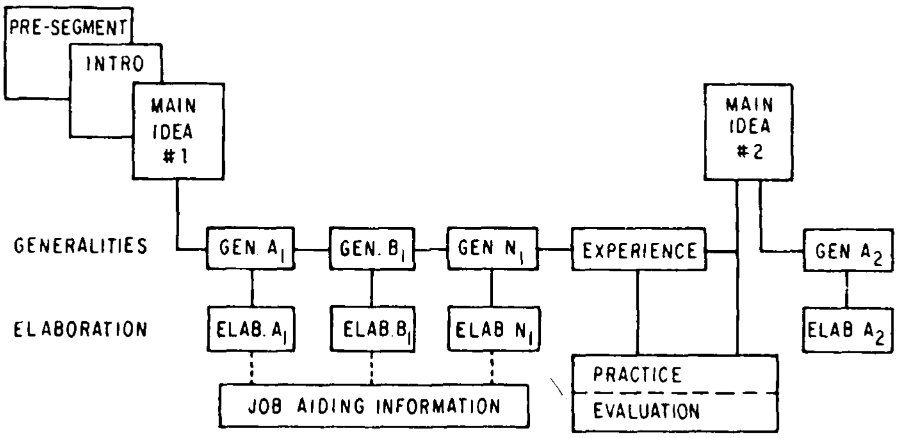


FIGURE 4 INSTRUCTIONAL MODEL FOR TPSS

HARDWARE/SOFTWARE

TPSS software requirements were met by vendor-supplied programs for the Job Entry Lessons and new programs developed by DRI to provide the Performance Support function. Commercially available hardware and software to satisfy traditional CAI and CMI requirements for development and delivery of instruction were identified through trade studies. The requirements addressed by the trade studies were: 1) hardware which utilized state-of-the-art microprocessor technology; 2) CAI software which contained "menu-driven" editors to simplify the development of instruction having high information density in text and graphic displays; and 3) CMI software with the capability to tailor a curriculum for each learner based on prospective program assignments. The ability to access external computer programs during delivery of Job Entry instruction was an additional software requirement. This capability permitted the independent development of the Performance Support software.

The hardware/software system chosen for the TPSS application is built by WICAT Systems, Inc. The hardware selected was configured to accommodate four persons working on lessons at the same time, and eight persons accessing the Performance Support job aids at the same time. The capacity to manage instruction via CMI was restricted to 100 persons enrolled at any given time in order to control the growth of management data files.

The Performance Support software capabilities developed to support the TPSS concept provide: 1) logon/logoff; 2) menu-driven selection of desired data bases available as instructional resources; 3) menu-driven access to the information data bases; 4) glossary requests; 5) assistance and help capabilities; and 6) mailbox capabilities. Additional software was developed, in the form of a device driver resident in the microcomputer, to make the Performance Support function available in different Program Office environments via the local area network being installed at Hanscom AFB.

CURRENT PLANS

Eleven event-oriented lessons, each prepared in five segments (approximately 150 displays per segment), will be used routinely in trials of TPSS during 1984. These lessons, together with their respective Performance Support information bases, and thirteen other single segment lessons (designed to treat other topics) represent the resources for use during the first year. Current Air Force plans also call for the installation of TPSS at the newly formed Acquisition Management Institute located at Brooks AFB, Texas.

AD-P003 316

Maintenance Troubleshooting Simulation: Adapting to the Trainee

J. Jeffrey Richardson, Ph.D.
Research Scientist
Denver Research Institute

Abstract

This paper presents a description of adaptive task centered maintenance simulation. First, the nature of this type of simulation is presented from the perspective of the trainee. Next, a method of on-line task analysis is presented through which adaptive task centered simulations are developed and delivered. Finally, the way in which the method of presentation, and even the structure, of a simulation can be adapted or altered dependent on evidence of student mastery is presented.

Introduction

Maintenance troubleshooting tasks certainly involve more than equipment manipulations. Behind each manipulation is a series of mental decisions which establishes its purpose. Therefore, a maintenance troubleshooting simulation must model not only equipment manipulations, but also the cognitive component of the task. Once such a model has been developed and used to generate a series of troubleshooting problems, it is possible to adapt the way the simulations interact with the trainee in ways that focus attention on unmastered components of the task, thus increasing training effectiveness and efficiency. For example, simulations often offer a menu of alternatives at each decision point. When a trainee demonstrates competence in selecting the correct option from the menu, the learning task may be made slightly more difficult by not displaying the menu. Now the trainee is expected to recall the response set from memory and then select the proper alternative. As a second example, when a complex procedure is simulated, task centered simulations require the trainee to explicitly state plans and goals leading to equipment actions before performing them. Once individual equipment actions are correctly performed, punctuated by explicit verbal statements of intent, the trainee may be allowed to "chain" together commonly co-occurring equipment actions into smooth sequences without intervening verbal mediation. This presentation will describe research on maintenance simulation funded by the Air Force Human Resources Laboratory which has produced a maintenance simulation system with the capability of adapting to trainee growth in competence in a way that is sensitive to differential rates of growth in competence in each subcomponent of the task. This adaptability is achieved through the use of an on-line task analysis which serves as a model of the trainee.

Interacting With a Task Centered Simulation

It will be helpful to the reader to understand task centered maintenance simulations from the perspective of a trainee. The trainee sits before a keyboard and two

video displays. One display represents the equipment. User interaction with this display is done through use of a touch panel. A second display represents the cognitive or problem solving component of the task. User interaction with this display is done through use of the keyboard.

The equipment display is a standard color video monitor on which a picture data base of the equipment relevant to the task is displayed. The monitor is equipped with a touch panel and, in a fashion that has become almost standard in microform and videodisc simulations of equipment, the student accesses controls, readouts, test points, and components by indicating with a touch of a finger on the video monitor face that portion of the picture to display in greater detail. For example, if a voltmeter range is to be set from 100 to 10, one begins by touching the voltmeter drawer on a rack of equipment pictured on the monitor. The rack picture is replaced with a picture of the voltmeter drawer. Next, the range control is touched and the voltmeter picture is replaced with a picture of the range control, currently positioned at 100. At this point, the "10" on the range scale is touched and the range control picture with the range control set to 10 is displayed. Now, the user might touch an area on the monitor screen labeled "done" to indicate that this equipment action is complete.

The task display is a medium or high resolution graphics display on which text, diagrams, decision alternatives, and feedback messages are presented. For example, suppose the trainee is developing a diagram of the signal path involved in a troubleshooting problem. The graphics display displays a skeletal block diagram, and a text message asks students on which part of the diagram they would like to elaborate next. At the bottom of the screen is a listing of the blocks from which to select a response. Given a correct choice, the graphics display now displays the block selected with the additional detail of the inputs and outputs. Now, the task is to identify for the situation at hand, specific signals for those inputs and outputs. Hence, a text message states: "Which signal would you like to add to the block diagram now?" At the bottom of the display, the decision alternatives to be used in answering this question are displayed. Continuing in this way, the task of detailing the signal flow would be completed.

A simulation proceeds from an initial task display in which the problem scenario is set, through a series of task displays leading to the first equipment action, through the equipment action simulated on the equipment display, and back to the next task display. This process continues, alternating between the task and equipment displays, until the problem is solved. The simulation is task centered in the sense that planning and problem solving activity controls access to equipment manipulation.

On-Line Task Analysis

On-line task analysis is used in order to develop and deliver adaptive troubleshooting simulations.

An on-line task analysis is defined as a computer-resident data base representing the set of goals and actions employed in accomplishing a task. It is a problem reduction tree (Barr & Feigenbaum, 1981-82). It represents a problem solving strategy known as hierarchical decomposition or planning by abstraction, wherein the troubleshooting problem is divided into simpler and simpler subproblems. This is done by repeatedly asking the question, "In order to get the job done, what must I do?", until the answer belongs to the assumed repertoire of the trainee.

The nature of the relationship between the task and the equipment is neatly represented by the on-line task analysis itself. The on-line task analysis has the form of a tree. The root node of the tree represents the task. From it emanate several branches to nodes representing the major subcomponents of the task. From each subcomponent emanate branches to nodes representing their decomposition, and so on, until the task is decomposed into elemental equipment actions. Therefore, all equipment actions are terminal nodes in the tree.

The target task of this research project was the operation and maintenance of an F-111 avionics test stand (Richardson, 1982; Pieper, 1982). The full analysis tree is up to 16 levels deep and contains between 1,000 and 2,000 nodes.

There are several benefits derived from the use of on-line task analysis. These include (1) the focus of the simulation is shifted from the equipment to the task, (2) computer-based simulation courseware is developed by simply flagging a path through the task analysis tree for each fault, (3) consistency in a problem solving approach is promoted at the same time allowing for flexibility in simulation design, and (4) interaction with the student at each simulation step can be adapted. These four issues are presented and discussed in detail in Richardson (in press). Of interest to the current paper is the fourth benefit.

Adapting to the Student

The way each step in a simulation can be individualized is by using a copy of the on-line task analysis as a student model. For each trainee, the on-line task analysis is annotated with a record of his or her performance at each node of the task analysis tree. Within a given simulation, and in the course of completing a series of simulations, a given node (representing a task component) will be encountered a number of times. A record of correct performance on each node is kept in the student model and used to modify the way in which the node is presented to the student. For nodes that do not involve equipment manipulation, a series of presentation modalities has been established: recognition, recall, performance, and summary.

Recognition means that the menu of selection options at the bottom of the display are in fact displayed. The trainee need only recognize the correct alternative from a set of choices, each of which is a legitimate choice in some situations, but only one of which is currently the correct choice. After the student has correctly performed this node a threshold number of times, it will shift into the next modality, recall.

In recall, the decision alternatives are not displayed, and the trainee must now recall the correct choice from memory, a more difficult learning task. Again, after responding correctly a threshold number of times, the modality of the node will change to performance.

With the performance modality, structural changes in the way in which the student interacts with the simulation occur. As stated above, a simulation proceeds by alternating between the task and the equipment displays. In general, each equipment manipulation is preceded by a task display in which the goal to perform the equipment manipulation is established. Trainees must "say" what it is they plan to do before they "do" it. But as trainees master a task, it becomes appropriate to support the chaining of commonly co-occurring equipment actions. That is, as a consequence of practice, a task becomes familiar and it is no longer appropriate to force trainees to plan separately for,

and then make, each elementary equipment manipulation. It is appropriate to allow the trainee to select the goal of performing a cluster of equipment manipulations and verbally stating each component manipulation within the cluster. With performance mode, the detail or graininess of the simulation adapts to the trainee's advancing level of competence. Again, this makes the learning task harder because the cognitive structure of the task, which is a mental crutch for remembering what to do, is removed. As competence progresses, nodes higher and higher in the on-line task representation shift to performance mode with the consequence that the simulation is more and more equipment centered rather than task centered.

The final modality is summary. Since it is the highest mode, by the time it is reached, the task has been performed under the most realistic and demanding setting within the context of simulation. The instructional assistance of the recognition and recall modes has been removed. The task centered nature of the task component has become equipment centered. Since the student has mastered this node, as evidenced by a number of correct trials as in performance mode, it becomes reasonable not to require the student to further practice this task component. Therefore, for a node in summary mode, the results of the completed task component are presented to the student in lieu of requiring the student to actually perform the task component.

The effect of this student modeling is to focus trainee attention on unmastered portions of the task, potentially promoting training effectiveness and reducing the time needed to complete a set of simulations. This is done on a completely individualized basis. The project is currently in its evaluation phase and the results of an experimental test of the training effectiveness of this approach is forthcoming.

References

Barr A., & Feigenbaum, E. (Eds.) The handbook of artificial intelligence (3 vols.). Los Altos, CA: William Kaufmann, 1981-1982.

Pieper, W. Graphics simulation system run time operation. Proceedings of the Association for the Development of Computer-Based Instructional Systems, (ADCIS), Vancouver, British Columbia, Canada, 1982.

Richardson, J. An integrated design and development system for graphics simulation. Proceedings of the Association for the Development of Computer-based Instructional Systems, (ADCIS), Vancouver, British Columbia, Canada, 1982.

Richardson, J. On-line task analyses in maintenance simulation. Machine-Mediated Learning, in press.

Acknowledgments

This work was supported by United States Air Force Human Resources Laboratory Contract No. *F33615-81-C-0006* with Essex Corporation, *William J.* Pieper, Principal Investigator. This work was performed as a collaborative effort between Denver Research Institute and Essex Corporation.

AD-P003 317

Evaluation of the Electronic Equipment Maintenance Trainer (EEMT): Individualized Troubleshooting Training[1]

Louis F. Cicchinelli, Ph.D.
Senior Research Psychologist
Denver Research Institute

Abstract

The two-dimensional Electronic Equipment Maintenance Trainer (2D EEMT) was designed and developed by the Navy as a generic trainer intended to overcome many of the problems associated with using operational equipment for training electronics technicians. The findings of the Training Capabilities Test (TCT) indicated that the 2D EEMT, when used in conjunction with operational equipment trainers, resulted in troubleshooting performance equivalent to that achieved when only operational trainers were used. Thus, the 2D EEMT can reduce reliance on operational equipment and increase the opportunity for individual student troubleshooting practice when used as a supplement to existing laboratory equipment. The 2D EEMT was found to be suitable for use in the Electronic Technician "A" School environment, and its generic nature makes it adaptable to other training environments with only changes in lessonware and the picture data base.

Introduction

Background. Operational readiness of naval electronic equipment depends, in large part, on the ability of enlisted technicians to perform appropriate maintenance actions. Effective technician training, which must be accomplished at reasonable cost, is then a necessary condition for military preparedness. Based on the assumption that maximum transfer of training will occur if technicians are trained on the same equipment that they are subsequently required to maintain, the Navy has traditionally used operational equipment to assist instructors in a variety of training tasks. In recent years, it has become clear that the use of operational equipment in training environments is not always the optimal approach to instruction. The use of operational equipment for training can be costly, dangerous, inflexible, unsupported by training materials, and, in some instances, the fleet is deprived of their full allotment of operational hardware. In addition, the basic premise that the transfer of training will be enhanced if the same equipment is used in the training and job environments is no longer as relevant. With the rapid changes in technology and the increased complexity and cost of electronic devices, more often than not, schools are forced to train students on outdated equipment which is itself difficult to maintain and no longer used in the fleet.

In response, the Navy initiated an engineering effort to develop a general purpose EEMT system that could reduce reliance on actual equipment trainers at the "A"

---

[1] This work was completed under contract with the Training Analysis and Evaluation Group, Naval Training Center, Orlando, FL. The views expressed in this paper are those of the author and do not necessarily reflect the view of the U.S. Department of the Navy.

School level, as well as provide an opportunity for increased troubleshooting practice during training. The EEMT system, and the 2D subsystem in particular, is a descendent of the Generalized Maintenance Training Simulator (GMTS). Field tests of the GMTS, a microfiche based system, have indicated that students could develop troubleshooting proficiency after practicing on the system and that students using the GMTS performed as well on final tests as students trained on actual equipment. While the hardware has changed substantially during this evolution from GMTS to EEMT, the functional characteristics of the two training systems remain essentially the same. The TCT of the EEMT system was designed to give equal emphasis to the evaluation of training effectiveness, suitability and cost of the 2D EEMT--an approach clearly supported by the Device Test and Evaluation Master Plan. The studies reported in this paper are part of the comprehensive TCT. The TCT was conducted in the Electronics Technician (ET) "A" School at Great Lakes Naval Station. All ET students receive training in Basic Electronics and Electricity, Advanced Electronics, and Communication followed by 9½ weeks of Radar Systems instruction. The primary application of the EEMT system was in this final course. In particular, 10 2D EEMT units were placed in the Primary Power (Area 1) laboratory section of Radar Systems instruction. This laboratory had 10 work stations using actual equipment trainers. An additional nine 2D EEMT units were placed in the extra study/remediation room associated with Radar Systems training.

EEMT system description. The functional training mission of the 2D EEMT was based on the premise that the training system would be generic. It was anticipated that more students could be trained in less time, more cost effectively, and to a greater proficiency level utilizing the EEMT system than with current actual equipment trainers. The 2D hardware consists of a microcomputer, instructor terminal, image and adaptive displays fitted with touch-sensitive panels for user input, and a videodisc player. The adaptive display presents text, problems, and menu choices; the image display presents equipment images from a photographic data base. This data base is arranged hierarchically and includes an overall system image together with successively more detailed images of system components and subcomponents. Using both structured and free-play lessonware, the 2D unit is primarily a stand-alone, interactive, videodisc-based trainer. The 2D EEMT system provided trainees with practice on generic preventive and corrective maintenance procedures through self-paced instruction, adaptable to the skills and needs of individual trainees. The system provided computer-assisted and limited computer-managed instruction such that student performance skills were tested and recorded. As implemented, only limited feedback on performance was provided.

Study 1: Analysis of Historical Performance Data

Purpose and procedure. Study 1 was designed to provide an initial assessment of the 2D EEMT's training effectiveness through a review of historical ET School student files. Student records were assigned to one of three independent groups as defined by the time periods of Area 1 AN/SPS-10 Radar training. These three time periods correspond to different levels of operational and 2D EEMT equipment availability for laboratory and extra study/remedial training. That is, students assigned to Group 1 received training on operational equipment only. For group 2 students, only operational equipment was available for laboratory training, and the 2D EEMT equipment was available for extra study/remediation. Group 3 students had access to both operational and 2D EEMT equipment for laboratory training, and the 2D EEMT was also available for extra study/remediation. Matched samples corresponding to the three time periods were created on the basis of student sex and ET School achievement prior to radar training. The resulting 3 samples of 72 students each were then divided into three equal subsamples (24 students each) on the basis of prior achievement to indicate students of low, middle, and high ability.

Results and discussion. The primary measure of student performance was the Performance Test administered individually to all students at the end of Area 1 Radar training. This test was a hands-on, timed, single fault troubleshooting test conducted using the operational AN/SPS-10 equipment.

Mean Performance Test scores as a function of training equipment availability were 83.92, 87.79, and 88.14 for Groups 1, 2, and 3 respectively. Mean scores increased from 79.88 for low ability students, to 86.12 for middle ability students, and to 91.75 for students of high ability. A 3x3 analysis of variance revealed significant main effects of both student ability [$F(2,207) = 10.05$, $p < .001$] and training equipment availability [$F(2,207) = 8.07$, $p < .001$]; the interaction between these variables was not significant. Thus, students trained during time periods when the 2D EEMT was available (Groups 2 and 3) performed better than those who had only conventional AN/SPS-10 training. Furthermore, this effect was apparently not dependent on the level of student ability as defined by prior ET School achievement.

However, since both EEMT Groups (2 and 3) performed equally well, it appeared that the locus of this effect was in the availability of the 2D EEMT for extra study/remedial training. Introduction of the EEMT into the Area 1 laboratory did not result in an additional performance increment. These findings also suggest that the EEMT can reduce reliance on operational prime equipment, although the extent to which the EEMTs were used for laboratory training (Group 3) was relatively slight. Finally, the results support the hypothesis that the 2D EEMT improves the performance of poor, moderate, and good learners. It must be noted that these findings are not conclusive since only equipment availability was used to establish the study groups; the precise training experiences of students were not known. Acquisition of the EEMT units was in part a response to the need for increased student flow to alleviate a student backlog. It is not clear to what extent this situation may have contributed to the higher performance scores seen during the Group 2 and Group 3 training periods. The next study in this section was conducted to provide confirmation of the Study 1 results.

Study 2: Training Effectiveness of the 2D EEMT

Purpose and procedure. The overall objective of this study was to determine the impact of incorporating the 2D EEMT into the existing ET School laboratory and extra study/remediation curricula. A total of 157 ET students were randomly assigned to one of six experimental conditions defined by crossing the three levels of training mode (high EEMT, low EEMT, and operational equipment only) with the two training formats (individual and pair). An analysis of variance (three levels of equipment training mode by two levels of training format) showed no significant differences among groups in prior ET School performance. Assignments were mixed within classes so that the effects of instructors, training shifts, and test problems were counterbalanced. The three training modes were defined by the extent to which 2D EEMT equipment was used to supplement a student's Area 1 laboratory and extra study/remediation training. Students in the operational equipment only conditions were not allowed to use the EEMT for any Area 1 training. Students in the high EEMT training groups worked an average of 3.90 troubleshooting problems on the 2D equipment and 1.18 during extra study. Students in the low EEMT groups worked an average of 1.98 problems in laboratory and another 1.15 during extra study. The average number of troubleshooting problems worked on the operational equipment during laboratory training for high EEMT, low EEMT, and operational equipment groups were 4.18, 6.53, and 7.55, respectively. In terms of operational equipment training time, these averages translate to 2.78, 4.59, and 6.24 hours for the three conditions, respectively. Students trained in an individual format worked all laboratory and extra study problems alone; those assigned to the pair format worked all

laboratory problems with another student but were allowed to attend extra study on their own.

Results and discussion. Performance test scores were the primary measure of training effectiveness used. Mean scores were 82.59, 78.63, and 80.11 for high EEMT, low EEMT, and operational only groups, respectively; 79.64 for individuals and 81.11 for pairs. A 3x2 analysis of variance conducted on these data revealed no significant main effect of either factor and no significant mode by format interaction. Mean time to completion were 31.70, 35.66, and 29.57 minutes respectively for high EEMT, low EEMT, and operational equipment only groups. Individually trained students took an average of 31.41 minutes and pair-trained students took an average of 34.19 minutes. 3x2 analysis of variance revealed a significant main effect of equipment mode [$F(2,141) = 3.76$, $p < .05$] and a marginally significant effect of training format [$F(1,141) = 3.41$, $p < .1$]. The mode by format interaction did not approach significance. Overall, then, while troubleshooting accuracy was not influenced by the equipment training mode or training format, troubleshooting efficiency was somewhat affected. Specifically, students trained on only operational equipment appeared more efficient than those who received 2D EEMT training, and students trained as individuals were slightly more efficient than those trained in pairs.

Unlike Study 1, this study found no appreciable differences among students' performance as a function of the training. The hypothesis that EEMT training improves performance was not supported. Differences in efficiency were observed, but they were easily attributed to familiarity with the Performance Test conditions; students who had more experience working on the operational equipment or working as individuals were faster in the individual operational equipment testing situation. The findings of Study 2 do suggest that the 2D EEMT can reduce reliance on operational equipment in ET training when used as a training supplement. Specifically, based on the high EEMT conditions, it appears that at least half of the operational equipment laboratory training can be replaced by 2D EEMT training without a decrement in student troubleshooting accuracy as long as EEMTs are also available for extra study/remediation purposes.

Study 3: Appropriateness of the 2D EEMT for ET "A" School Training

Purpose and procedure. This study was designed to evaluate the compatibility of the 2D EEMT with the training environment and objectives, attention to safety issues, person-equipment interfaces and lessonware content. The information necessary to evaluate these issues was obtained from structured interviews with 48 ET "A" School instructors and supervisors familiar with the Area 1 AN/SPS-10 training program. Additionally, student critiques were completed by 109 students following their Area 1 laboratory training. All respondents were asked to use a five point scale to rate their agreements with statements about the operational equipment training and, if applicable, about the 2D EEMT training they received. Mean ratings were calculated and are shown in parentheses. Two-tailed t-tests were then computed for each statement to identify those mean ratings that differed significantly (at the .05 level) from a noncommital rating (i.e., 3.00).

Results and discussion. In responding to statements that focused specifically on the free play troubleshooting lessonware, the authoring software, and the management aspects of the 2D EEMT system, instructors disagreed with statements that: on a class basis, EEMT system management does not add significantly to the workload (2.42), student data produced by the EEMT system are helpful in student management (2.42), student data produced by the EEMT system are useful in assessing student abilities (2.36), and the authoring software provides an effective method of developing new lessons (2.39). It

should be noted, however, that only five of the respondents indicated having hands-on experience with the authoring software.

Of the 34 respondents who provided written comments on various aspects of the 2D EEMT system, 11 instructors believed there was a general need to improve system reliability and 9 more noted specific reliability problems (e.g., touch panel drift, system lock-up) that interfered with training. Another major concern was the inability of the system to provide appropriate safety training and/or monitor safety-related student errors. Additional comments focused on the need for better representations of test equipment use, improved data management, and expansion of the photographic data base. A frequent remark was that the EEMT is good for remedial/extra study applications but it cannot take the place of the actual equipment for laboratory training.

Overall, student responses to 11 statements that focused on EEMT training were either noncommital or marginally favorable. Students believed that there was a variety of lessons provided (3.49), the equipment was easy to use (3.50) and comfortable (3.36), and that training time was mostly spent in troubleshooting (3.53). Students disagreed with statements that the lessons were too long (2.43) and that training time on the equipment was often wasted (2.47). The most common student likes about the 2D EEMT training were ease of use, comfort, reduced safety risk, availability for practice troubleshooting, and ability to solve more problems in a given period of time. The most common student dislikes were equipment malfunctions (particularly touch panel drift and EEMT lock-up), use of test equipment, inability to access components properly (or read reference designations), and the possibility of developing bad safety habits. Nineteen students remarked that the EEMT is a good training aid but it should not replace the operational equipment; 13 students thought the EEMT should be used for remediation and extra study; six students preferred more operational equipment training.

When comparing the 2D EEMT with operational equipment trainers, students thought that training time was wasted more on the EEMT (2.47) than on the operational (1.67) equipment [$T(95) = 5.40$, $p < .001$] and that equipment breakdowns interfered more with training on the EEMT (1.60) than on the operational (2.49) equipment [$T(94) = 5.91$, $p < .001$]. With respect to the variety of lessons provided, however, the EEMT (3.94) was rated slightly higher than the operational (3.49) equipment training [$T(88) = 3.18$, $p < .01$].

General Conclusions

The 2D EEMT, a videodisc-based system, is an effective supplement to operational equipment trainers in a hands-on instructional setting. It is particularly appropriate for remediation/extra study in that it provides a relatively safe training environment requiring little supervision. Some hardware modification is necessary to improve reliability and the ease of operating the system. Although it is possible to group lessons into pools based on difficulty, as currently configured, student progress and errors in solving specific troubleshooting problems are not used to direct further instruction. Changes in the software can be expected to make the 2D EEMT more responsive to individual student instructional requirements.

Factors Influencing the Selection of Individualized Instructional Approaches

Anita S. West, Ph.D.
Division Head, Senior Research Mathematician
Denver Research Institute

Abstract

Recent research conducted by DRI for the Air Force Human Resources Laboratory has focused on the selection and implementation of self-paced resident training methodologies, including factors that influence the selection of computer-based technologies and factors impacting the successful implementation of nonconventional instructional roles. The results of these research programs are consolidated and presented in the form of a "first cut" decision aid for training managers faced with the decision of whether or not to individualize their resident training programs.

Introduction

Training and learning research at the Denver Research Institute has been focused, during the last few years, on various aspects of individualized instruction (I.I.) in military training, from the evaluation of the use of simulators and applications of artificial intelligence for teaching troubleshooting skills to the examination of factors that influence the success of self-paced training and the development of decision aids for managers considering the use of CAI for a given course or group of courses. Further, DRI personnel have been investigating the role of the training instructor as it has evolved as a function of the use of various nonconventional teaching formats. Through these research studies, the advantages and disadvantages of various forms of individualized instruction have been investigated. In many respects, each application and each training delivery mechanism is different from the others and there are different factors that suggest the match of each. However, some general kinds of guidelines about group vs. individual instruction have emerged from these studies. This paper is an attempt to consolidate some of these findings in a useful way for managers who have the opportunity and the responsibility for recommending and selecting training delivery options. The paper first briefly describes three research studies recently conducted at DRI and then translates their collective conclusions through a "first cut" decision aid for those considering an individualized training format.

Factors Influencing the Success of Self-Paced Instruction in Air Force Technical Training

During the summer of 1982, DRI was funded by the Air Force Human Resources Laboratory to take a critical look at self-pacing in resident Air Force training. The concern was expressed that, in spite of a literature that produced a wealth of support for the concept of self-paced instruction, the Air Force was experiencing a considerable amount of dissatisfaction with it and, in fact, the number of Air Training Command courses that were currently self-paced had dwindled from a high of 110 in 1979 to only about 25 in 1982. The investigation was conducted as a rather intensive 4-month study at a level of effort of almost one person year using a case study approach. The full report on the investigation written by McCombs, Back and West, is expected to be available as a technical paper in early 1984.

The study began with an exhaustive noncritical study of the literature to develop an inclusive list of hypothesized factors. After organizing the factors into a conceptual model to guide the data collection, six courses were selected that met study-developed success criteria and another six courses were selected that were generally agreed to be unsuccessful in their self-paced formats. An intensive case study analysis of each of the 12 courses was conducted through interviews, observations and review of archival data. The study then merged the three sets of case study data with the literature model. The data sources included (1) hard data on the course and the students, e.g., student flow, student aptitute scores, course history, how long self-paced, etc., (2) observational data, e.g., teaching methods and environment, study environment, presentation quality of materials, etc., and (3) attitudes and perceptions of training personnel (through interviews) and of field and student personnel (through written evaluations).

Results. An objective of the study had been to identify several individual factors that one could point to and say "these specific characteristics or activities made the difference" where those factors would be in addition to the factors normally associated with the success or failure of any course. The study, however, resulted in a much more complicated analysis where no single factor stood out as both a necessary and sufficient condition for success. A number of factors, however, were consistently associated with success and a number of others with failure and these, for the most part, were the mirror images of one another. For example, high instructor motivation was associated with success and low instructor motivation with failure. These two-sided observations were interpreted to give additional reliability to the observations; however, it was not observed in all cases, e.g., low reading ability was associated with failure but high reading ability did not surface as a success criterion. What the study was able to produce was a somewhat complicated relationship of factors grouped under the headings of Management and Instructional Issues. An illustration and description of these factors appear in the full report. Very briefly, it appears that self-pacing was effective (or at least not ineffective) in almost every case. An exception was the course that required frequent team performance. However, in order to overcome the administrative costs and personnel efforts required to put self-pacing in place and maintain it and the initial student and instructor resentment frequently observed, there had to be a sense of "benefits gained" or "value added" to make the efforts worthwhile. This generally meant perceptions of cost and time effectiveness on the parts of both training personnel and students, and real opportunities for instructor control and flexibility.

Factors Influencing the Decision to Adopt CAI

In response to a perceived need among mid-level training managers to learn more about computer-assisted instruction and various CAI options and to receive assistance with the decision process, DRI researched and produced "The CAI Decision Handbook," (Kemner-Richardson, Lamos and West, AFHRL, in press). The handbook was written for two applications: (1) as a resource and reference guide for Air Force instructional managers considering the adoption or expansion of CAI at some future time and (2) as a decision aid for an instructional manager currently involved in the CAI decision process. The use of the handbook is appropriate when considering the implementation of CAI in a new course or in an existing course. Its contents are designed to be helpful where CAI systems are already available to the trainers as well as in those cases where the adoption of CAI requires new system acquisition.

Some assumptions were made by the authors regarding the target audience and their needs. In general, it was assumed that readers possessed a significant amount of knowledge about training and about Air Force instructional procedures, but had minimum experience with computer-assisted instruction. One major section of the handbook

provides basic information about various hardware, software and courseware options. Entitled, "Critical Factors for Adopting CAI," it is a body of reference information on many of the major factors, including relative costs, that need to be considered before coming to even a preliminary decision about the use of CAI. In addition, a portion of the handbook presents specific tools, in the form of worksheets, for evaluating the need for CAI, for identifying configurations most closely matched to instructional needs and practices, and for estimating the feasibility of initiating CAI implementation. It provides a series of questions and a flow chart to help organize the planning and decision process. A final section discusses some nontechnical issues of concern when introducing any kind of change into an existing organization.

Results. The factors that suggest the use of individualized instruction in general and CAI in particular fall under two major headings: Organizational Needs (critical training throughput requirements, shortage of qualified and trained instructors, course development/revision time available, the need to standardize content and procedures, etc.) and Instructional Needs (course is procedural and practice intensive, students vary widely in background and previous training, the course is currently self-paced and student reading ability is inadequate for programmed text, etc.). After a determination that CAI is indicated, a procedure for estimating the costs of appropriate CAI options is recommended to ensure that all of the resources will be available and that the benefits anticipated exceed the anticipated costs. A final caveat is provided that explains the decision to adopt CAI or forgo adoption of the medium cannot be made explicitly prescriptive. The decision-aiding method merely highlights the many relevant factors and issues.

Personnel Roles and Requirements for Nonconventional Instruction

That instructor role definition is related to the overall success of nonconventional instruction has been established (McCombs, Back & West, 1984) and these findings are consistent with the general literature (Back & McCombs, 1984). A theoretical analysis performed by McCombs and Dobrovolny (1980) specified seven theoretically ideal roles for instructors in a computer-managed environment. More recently, DRI investigated the issue of whether the theoretical model developed for CMI environments was generalizable to other nonconventional instructional (NCI) formats found in Air Force technical training where each of the NCI environments can be characterized as individualized instruction. The study also sought to identify the extent to which these roles are invariant as a function of NCI format, the problems experienced by instructors in performing their NCI roles and whether or not the roles and behaviors currently performed by instructors could be more efficiently performed by non-instructor personnel. Finally, the study looked for implications for the design of an instructor training curriculum.

Since an accurate and adequate definition of instructor roles is necessary for determining appropriate levels of staffing and instructor training, as well as being related to the effectiveness of the instruction and the success of the individualized instructional format, the results of this study represent an important component in the examination of factors that influence the selection of individualized instructional formats.

Results. Although absolute percentages of time in each role varied as a function of each NCI format investigated, the analysis by McCombs and Lockhart indicated that the ideal role order of counselor, modelor, evaluator, diagnostician, remediator, implementor and planner was judged to be an appropriate model for the NCI instructor. Study findings also indicated that NCI instructors spend almost 20 percent of their time

performing roles that were not included in the theoretical model, such as author, equipment maintainer, supervisor, and a miscellaneous category that consisted of a variety of clerical and administrative activities. The complete findings and their implications for instructor training are available in the final technical report (McCombs & Lockhart, 1983).

Conclusions

Based on these and other studies, it can be seen that the preliminary selection of an individualized instructional format, whether it be in a computer-based or manual self-pacing environment, can be rationally determined based on several factors relating to the training mission and course objectives, the training management environment, and instructor preparation and orientation. The two tables presented below serve to (1) identify training needs that suggest for or against the use of individualized instruction and (2) present the factors required for a successful implementation.

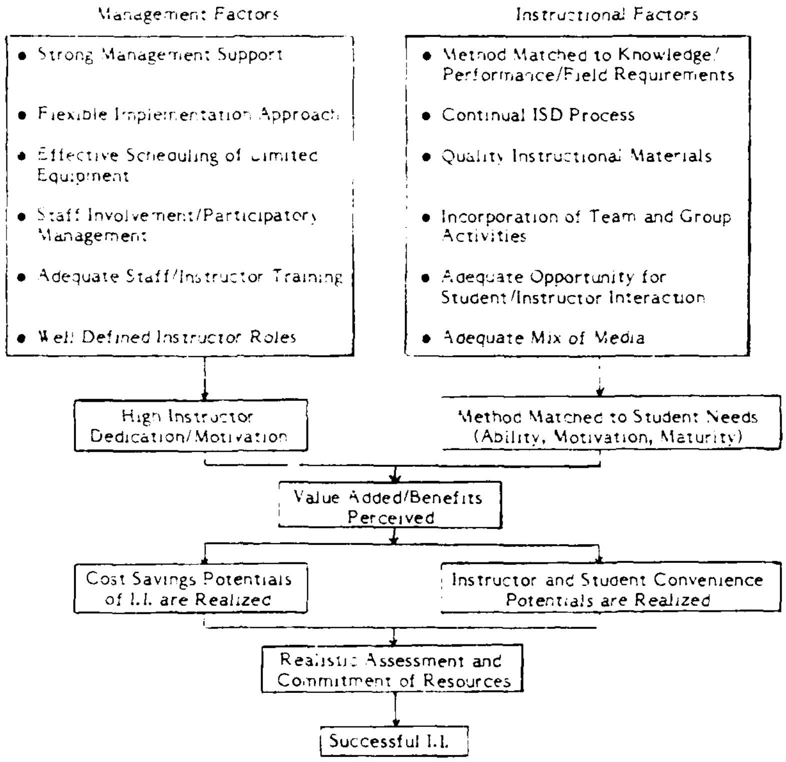
Table 1

Indications For and Against Individualized Instruction

I.	Training Objectives that Suggest Individualized Instruction
	• pretrain in basic skills as a prerequisite to teaching required skills • reduce student waiting time, especially where Trained Personnel Requirements are sporadic • accommodate highly variable aptitude/experience levels among students • reduce requirements for highly experienced/field experienced training staff • reduce amount of time away from job for training • standardize course content and procedures from class to class • reuse blocks of information that are non-AFSC specific in other classes • provide training in a course that is procedural and practice intensive
II.	Training Objectives/Conditions that Contraindicate Individualized Instruction
	• provide familiarity with dangerous or infrequently operating conditions that cannot be simulated • provide a great deal of hands-on experience that cannot be simulated • train personnel to work predominantly as part of an integrated team • provide a substantial degree of interpersonal skills for adequate job performance • group-paced format is currently satisfying all administrative and instructional objectives

Table 2

Factors Required for Successful Implementation



References

Back, S. M., & McCombs, B. L. Factors related to the success or failure of self-pac instruction: A literature review. (To be published as a technical paper by AFHRL, Low Air Force Base, CO, 1984). Denver, CO: Denver Research Institute, 1983.

Kemner-Richardson, S., Lamos, J., & West, A. S. The CAI decision handbook. (To published by AFHRL, Lowry Air Force Base, CO, 1984). Denver, CO: Denver Resear Institute, 1983.

McCombs, B. L., & Dobrovolny, J. L. Theoretical definition of instructor role in comput managed instruction. (NPRDC-TN-80-10). San Diego, CA: Navy Personnel Research Development Center, March 1980.

McCombs, B. L. Personnel roles and requirements for non-conventional instruction (NC Paper presented at the Military Testing Association, Pensacola, FL, October 1983.

McCombs, B. L., Back, S., and West, A. S. Factors critical to the success or failure of se paced instruction in Air Force technical training. (To be published as a technical paper AFHRL, Lowry Air Force Base, CO, 1984). Denver, CO: Denver Research Institu 1983.

S., & McCombs, B. L. Factors critical to the success of self-paced instructi presented at the Military Testing Association, Pensacola, FL, October 1983.

The Structure of Processing Resource Demands in Monitoring Automatic Systems

John Micalizzi

Ballistic Missile Office, Norton AFB, California

Abstract

Human operators are increasingly being called upon to function as monitors of automatic systems. System monitors, as opposed to active controllers, do not necessarily experience lower workload levels during task performance. In fact, prior research has suggested that workload demands may not be reduced but rather shifted to a functionally separate processing "pool" according to a structure specific view of human attention. Sternberg's additive factors method may provide a useful workload assessment technique for localizing the information processing demands of task performance. The present study couples a primary failure detection task with a secondary Sternberg task which employed a perceptual and response load manipulation. The results demonstrated a significant overlap of processing resources for the failure detection task and the Sternberg perceptual condition. For the response load condition, there was no evidence of shared resources between the two tasks. These results have significant implications for task configuration and workload assessment research.

Introduction

Background

The role of the human operator has undergone dramatic revision in recent years with the continued encroachment of automation into the human domain of man/machine systems. The human operator's function has steadily evolved from that of a manual controller of dynamic systems to the role of a monitor and supervisor of automated systems. Unfortunately, this evolution, in many cases, has not been accompanied by a corresponding development of sensitive research methods to investigate the qualitatively different demands which are placed on human monitors. Automation will never eliminate operator effort and its introduction into the task situation may not necessarily serve to reduce the load experienced by human monitors. The purpose of this study is to further explore the information processing demands of monitoring automatic systems.

The Concept of Processing Resources

Workload assessment research represents a variety of concepts, models and methodologies which attempt to quantify the demand placed on the operator's limited information processing abilities. The concept of selective attention has been developed in an effort to account for the processing limitations of human performance. While structural theorists (e.g., Keele, 1973) prefer to view processing resources as related to the

discrete competition of tasks for specific processing mechanisms, proponents of capacity theories (e.g., Kahneman, 1973) emphasize the flexible nature of processing resources which permits allocation in response to task demands.

The structure specific resource model represents a compromise between the structural and capacity views of human attention. The notion of a number of separate processing reservoirs (as opposed to an undifferentiated pool) is consistent with many results reported in the dual task literature. Wickens(1980) has reviewed the results of many dual task studies and his efforts have led to several promising candidates for resource definition including: stages of processing (perceptual-central processing-response), modalities on input (visual-auditory) and output (vocal-manual), and hemispheres of processing (verbal-spatial). This multiple reservoir view envisions task interference as a function of processing pool overlap.

The secondary task methodology for workload assessment is a useful technique for investigating the information processing demands involved in task performance. If a capacity model of processing resources is adopted, then workload may be conceptualized as the proportion of total resources demanded by a particular task. The higher the workload of a primary task, the less "residual capacity" is left available for performing any concurrent secondary task. Decrements in secondary task performance can then serve as an index of primary task demands (Ogden, Levine, & Eisner, 1979).

Sternberg's Additive Factors Method

The additive factors method is based on Sternberg's (1967) investigations into the scanning of human memory. The approach taken by Sternberg assumes that the reaction time interval is filled with a sequence of independent stages of processing. Total reaction time, then, is simply the sum of the individual stage durations. When an experimental manipulation (factor) affects reaction time for a particular information processing task, it changes the duration of one or more of the constituent stages of processing. If two experimental manipulations affect two different stages, they will produce additive effects on total reaction time. However, if two experimental factors interact, so that the effect of one factor is dependent on the level of the other, they must affect some stage in common.

Workload Applications

The additive factors method has displayed an encouraging potential as a methodology for assessing primary task workload (Spicuzza, Pincus & O'Donnell, 1974; Crawford, Pearson & Hoffman, 1978). Variations of the Sternberg paradigm have coupled a monitoring task with the central processing manipulation to determine the locus of monitoring resource demands (Wickens & Micalizzi, 1980). Preliminary results are inconclusive in establishing the central processing reservoir as the source of monitoring processing load.

The present investigation represents a follow-on to this monitoring study, requiring subjects to passively monitor a dynamic system and detect

failures while performing a concurrent Sternberg secondary task. The quantitative demands of failure detection will be varied across subjects by manipulating the cutoff frequency of the random noise function. The Sternberg manipulations will load the perceptual and response processing stages. Prior research has suggested that this failure detection task demands primarily perceptual processing resources and, thus, should interact with the perceptual Sternberg manipulation. These results would suggest that Sternberg's additive factors method would provide a very efficient and effective tool for exploring the multidimensionality of workload demands.

Method

Subjects

Eight right handed male undergraduate students from the University of Illinois volunteered to participate in all experimental manipulations. All subjects had normal vision and were paid $2.50 per hour plus additional bonuses.

Tasks

Failure detection. This task is similar to the automatic mode (AU) of failure detection reported in Wickens and Kessel (1979). In the present study, subjects were required to monitor a single axis pursuit tracking display which moved horizontally across the Cathode Ray Tube (CRT). System failures were simulated by a 10 second linear ramp change in dynamics from a first to second order system. The computer recorded hit latency and false alarms.

Sternberg task. The general Sternberg paradigm required subjects to recognize previously presented spatial information consisting of randomized dot patterns. Subjects responded by pressing the appropriate key to indicate whether the test stimulus matched the memorized stimulus. A perceptually loaded condition and a response loaded condition were utilized to draw additional attentional resources from these stages of processing.

Experimental design

A within subject design was employed where each subject participated in all experimental manipulations. The various Sternberg conditions included a baseline condition, a perceptually loaded condition and a response loaded condition. The failure detection difficulty manipulation varied the cutoff frequency of the random noise function from .32 Hz to .5 Hz. Each of these task manipulations was performed under both single and dual task conditions.

Results/Discussion

The experimental results indicate that the interaction between perceptual load and the presence of the failure detection task was statistically significant, $F(2,14) = 8.10$, $p < .01$. Under dual task conditions, a signi-

ficantly greater increase in Sternberg reaction times was obtained for the mask manipulation compared to the no mask condition. In constrast, no significant positive interaction was found between response load and failure detection. The double key response reaction times were not as severely disrupted under dual task demands as the perceptual load reaction times.

The ability of subjects to maintain consistent prima.y task performance for both single and dual task conditions is an important requirement for any interpretation of the dual task data. A comparison of the single and dual ask failure detection data revealed essentially equivalent performance for these two conditions, $F(3,21) = 2.17$, $p > .122$. In most cases, subjects were able to maintain superior performance under dual task demands. Thus, we can be relatively secure in the knowledge that similar amounts of processing resources for the failure detection task were used in single as well as dual task conditions. This assumption permits an interpretation of Sternberg reaction time decrements as an indication of task manipulations.

Although maintaining single task failure detection performance under dual task demands is an important requirment for any interpretation of the reaction time data, an equally important consideration is the ability of subjects to avoid utilizing a "resource tradeoff" strategy in producing the observed reaction time decrements.

Under this interpretation, processing resources are assumed to be diverted (traded off) from the Sternberg task (resulting in higher reaction times) and applied to the failure detection task (resulting in lower hit latencies). As a result, variations in reaction time performance across Sternberg conditions could be explained in terms of subject strategy without reference to competition among hypothesized pools of processing resources. A comparison of mean dual task failure detection latencies across Sternberg conditions reveals no significant variations, $F(2,14) = .234$, $p > .794$. The relatively higher perceptual load reaction times observed under dual task conditions were not necessarily accompanied by correspondingly lower hit latencies. The variation of dual task failure detection performance is not large enough to account for the larger decrements in reaction time performance.

Similar analyses of the error data for both the Sternberg and failure detection tasks also failed to reveal a speed/accuracy tradeoff strategy.

The experimental results provide at least some support for the main hypotheses advanced in the beginning of this paper. First, the significant interaction between perceptual load and failure detection demands indicates some degree of processing resource overlap between these two tasks within the framework of the additive factors method. Second, the lack of a significant interaction between response load and failure detection demands provides evidence for the notion of a separation of the respective processing resource pools.

The failure detection task used in this study appears to be primarily perceptually loaded. This conclusion is consistent with previous studies

(Wickens & Kessel, 1979) which investigated the resource demands of failure detection with a different secondary task. In addition, these results also support a stages of processing dimension for the structure specific resource model which is particularly applicable to workload investigations. The general Sternberg paradigm utilized in this study has shown promise as a technique for probing the multidimensionality of workload demands within the context of dual task methodology.

References

Crawford, B.M., Pearson, W.H., & Hoffman, M.S. Multipurpose switching and flight control workload. Proceedings of the Sixth Symposium on Psychology in the Department of Defense, 1978, 219-221.

Kahneman, D. Attention and effort. Englewood Cliffs, New Jersey: Prentice Hall, 1973.

Keele, S.W. Attention and human performance. Pacific Palisades, CA.: Goodyear Publishing Co., 1973.

Ogden, G.D., Levine, J.M., & Eisner, E.J. Measurement of workload by secondary tasks. Human Factors, 1979, 21(5), 529-548.

Spicuzza, R., Picus, A., & O'Donnell, R.D. Development of performance assessment methodology for the digital avionics information system. Dayton, Ohio: Systems Research Laboratories, Inc., August 1974.

Sternberg, S. Two operations in character recognition: Some evidence from reaction time measurements. Perception and Psychophysics, 1967, 2, 45-53.

Wickens, C.D. The structure of attentional resources. In R. Nickerson & R. Pew (Eds.). Attention and Performance VIII. Englewood Cliffs, New Jersey: Lawrence Erlbaum Associates, 1980.

Wickens, C.D., & Kessel, C.J. The effects of participatory mode and task workload on the detection of dynamic system failures. IEEE Transactions on Systems, Man and Cybernetics, 1979, SMC-13, 24-31.

Wickens, C.D., & Kessel, C.J. The processing resource demands of failure detection in dynamic systems (Tech. Rep. EPL-79-1/AFOSR-79-1). Champaign-Urbana, Illinois: University of Illinois, Engineering-Psychology Laboratory, January 1979.

Wickens, C. D., Sandry, D., & Micalizzi, J. A validation of the spatial variant of the Sternberg Memory Search Task. University of Illinois Engineering Psychology Laboratory Technical Report EPL-81-2/ONR-81-2, June 1981.

AD-P003 320

MAGIC 1: A Study of Manual Versus Vocal Control Under Heavy Task Loading

David G. Curry
Flight Dynamics Laboratory
Wright Patterson Air Force Base, Ohio

Abstract

The Microcomputer Applications of Graphics and Interactive Communications study, Phase I (MAGIC 1) was designed to compare the effectiveness of a voice activated control system with that of a multifunction control panel utilizing a branching control logic under conditions of heavy manual-visual task loading. Eighteen Air Force personnel were asked to perform fourteen different aircraft subsystem control tasks using each of the control modes while simultaneously performing a complex loading task. Results showed that the subjects demonstrated significantly better performance on three of the five dependent variables when using the voice system, indicating that they were better able to maintain concentration on the loading task with this mode. Analysis of subjective data showed an overwhelming subject preference for the voice control mode.

Introduction

One of the major problems facing both designers and pilots of high performance aircraft is the explosion of instruments and controls which must be incorporated into the cockpit. We have progressed from the French Spad of World War One with its eleven controls and displays to the current F-15 with more than 300. Clearly, this presents dificulties in simply finding room to physically locate a particular device within an extremely limited area; when the problems of both rapidly locating and then accessing the device are considered, the problem becomes enormous.

One method of addressing this situation is through the use of a speech recognition system to provide the pilot with the option of using voice as a control mode. Voice presents several advantages. First, speech is man's most natural mode of interaction, one we are exposed to from birth; no new skills need be learned. Second, such a system allows the pilot to keep his hands constantly on the primary flight controls; there is no interruption in the smooth operation of the aircraft. Third, it allows the pilot to keep his attention focused outside of the cockpit, since he does not have to look inside the cockpit to locate a particular control. Finally, the use of voice as a second control mode allows the pilot to operate two separate controls simultaneously if necessary (one manual, one by voice).

Method

Apparatus: The dynamic mock-up used in this study was a single-seat cockpit configured to F-15 dimensions (See Fig 1). The cockpit was fitted with five color cathode ray tubes; the head-up display provided the subject with the loading task, while the central CRT served as a multi-function control unit on which the manual control operations were performed. A separate digit entry keyboard was located on the lower left portion of the instrument panel. The other three screens displayed a stick map, engine

status, and stores status, none of which were critical to the current study. The speech recognition system used was a Votan V5000 (an isolated-word, speaker-dependent system). The loading task was generated by a Colecovision video arcade system equipped with a Cosmic Avenger cartridge. This task was selected because it provided both high task loading and a motivation source for the subjects.

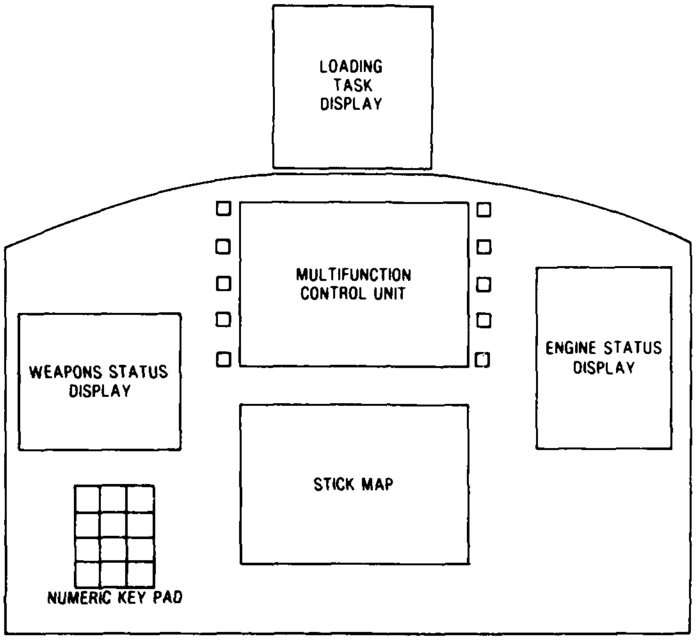


Figure 1. Mockup Configuration

Subjects: The subjects selected for this study were eighteen Air Force personnel possessed of prior video gaming experience. Potential subjects were screened by being given one-half hour worth of training on the loading task and then having their performance at the end of this period compared to a preset minimum acceptable value.

Procedure: Subjects were first given an extensive initial briefing on the cockpit, the loading task to be used in the study, and the types of subsystem control tasks which would be performed. Following this they were given a one-half hour training/practice/screening period using the loading task. Minimum acceptable performance on the loading task was defined as an average of 30 seconds survival time for each ship of a five-ship mission. If the subject met this minimum criterion, he was then trained in the particular control mode being evaluated first. For the voice control mode, this consisted of first loading the speech recognizer with the subject's voice patterns for each of the words in the 55-word command vocabulary, then having the experimenter walk the subject through two "dry-runs" of a similar series of control tasks without the loading task, and then the subject operating the system by himself with the loading task operational during a practice run. For the manual condition, the subject was shown the particular series of switch hits required for each of the control operations on an initial training pass performed by the experimenter; next, the subject operated the controls himself on a second pass under experimenter instruction, and, finally, a

training pass with the loading task and control tasks both operational was completed. Data collection runs for each mode were performed immediately following the final training pass. A post-mission questionnaire was given to the subjects following the final data mission.

Tasks used in this study consisted of 13 standard control operations using the communications, navigation, and stores management systems. These tasks were vocally initiated by the experimenter in a randomized order. Relative to the experimental design, the order of both the tasks and the control modes were counter balanced by means of a Latin Squares technique to eliminate any systematic order effects. Performance measures used in this study included: 1) the average score per ship obtained on the loading task, 2) the number of times one of the subject's ships was destroyed during a subsystem control task, 3) the amount of time required for the subject to initiate a subsystem task after being commanded by the experimenter, 4) the number of control operation errors on the subsystem tasks, and 5) the time required to complete the subsystem tasks.

Results

Control operation data and loading task performance data were analyzed by multivariate analysis of variance (MANOVA) using the Statistical Package for the Social Sciences, Version 8.3 program (Nie, Hull, Jenkins, Steinbrenner, and Bent, 1975) and significant effects are summarized in Figures 2, 3, and 4. A significant main effect was found for the input mode, $F(5,13)=16.16$, $p<.00003$. A Finite Intersection Test (FIT), a simultaneous comparison test for multivariate data, was utilized to determine which of the dependent variables were most responsive to changes in the independent variable (Cox, Krishnaiah, Lee, Reising, and Schuurman, 1980). According to the FIT, the number of the subject's ships which were destroyed while operating subsystem controls was significantly lower for the voice mode than for the manual, $F(1,34)=16.692$, $p<.01$; the same was true of both the time required to initia e a task and the time required to complete the tasks; $F(1,33)=13.307$, $p<.01$ and $F(1,31)=24.436$, $p<.01$, respectively.

Subjective data from the post-experimental questionnaires was analyzed by use of Kolmogorov-Smirnov one-sample tests (Kirk, 1968). Results showed that the subjects found both the voice and manual modes to be from moderately to very easy to use ($\underline{D}(18)=.544$, $p<.01$, and $D(18)=.322$, $p<.05$, respectively), and felt that the voice mode would prove to be a great advantage in maintaining heads up flying ($\underline{D}(18)=.800$, $p<.01$). When given a choice between the manual and voice control modes, voice was universally preferred ($\underline{D}(18)=.633$, $p<.01$).

Discussion

The use of the voice control mode was seen to significantly reduce the number of kills resulting from the subsystem control tasks; from this it can be inferred that the subjects could concentrate substantially more of their attention on the loading task. In an actual aircraft, this would translate to a greater ability to maintain head-up flying, a desireable outcome at any time, and considerably more so in a hostile environment. The decrease in both initiation and operation time also provides obvious advantages for time critical tasks. The lack of significance on the part of the error and score variables is probably accounted for by the high correlation of the former with the kills metric (.5489), and the high variance of the latter (353,433).

Subjective data regarding the ease of use of the manual multifunction control exhibited no significant clustering effect ($p>.05$), with the subject's responses ranging about equally between "somewhat difficult", "moderately simple", and "very simple." When the same question was asked for the voice mode, responses were significantly clustered around the "moderately easy" and "very easy" levels ($p<.01$), with only one subject assigning the system a lower rating. All of the subjects preferred the voice entry mode to the manual, regardless of their performance under each condition, suggesting that the perceived workload under the voice condition was sigificantly lower.

This agreement between the objective data supporting the voice control mode and the subject's significant preference for it combine to suggest that such a control mode would be a highly viable option for incorporation into the cockpits of future high performance aircraft and, indeed, in any other environment where space is critical and strict attention to a primary task is of paramount importance. Other application areas might include nuclear reactor control rooms and airport control towers.

References

Cox, C.M.; Krishnaiah, P.R.; Lee, J.C.; Reising, J.M.; and Schuurman, F.J. A study on Finite Intersection Test for multiple comparisons of means. In P.R. Krishnaiah (Ed.) Multivariate Analysis (Vol. V). Amsterdam: North-Holland Publishing Company, 1980.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing Company, 1968.

Nie, N.H.; Hull, C.H.; Jenkins, J.G.; Steinbrenner, K.; and Bent, D.H. Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill Book Company, 1975.

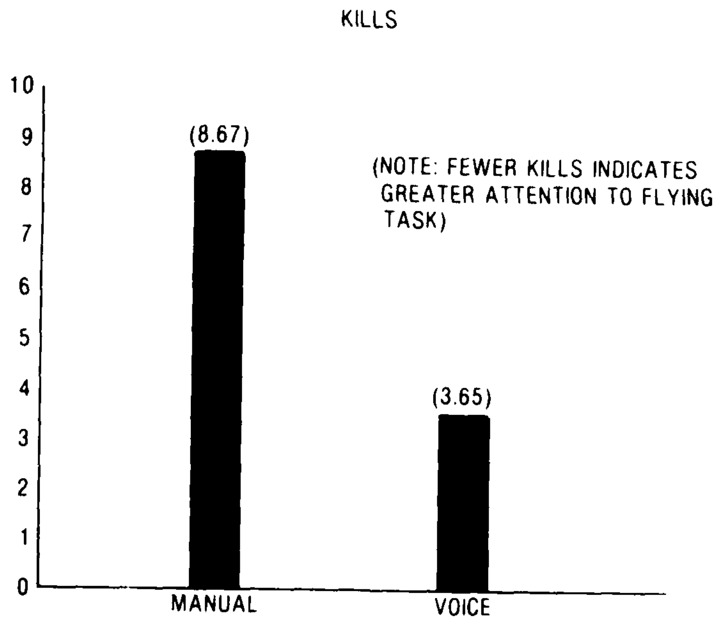


Figure 2

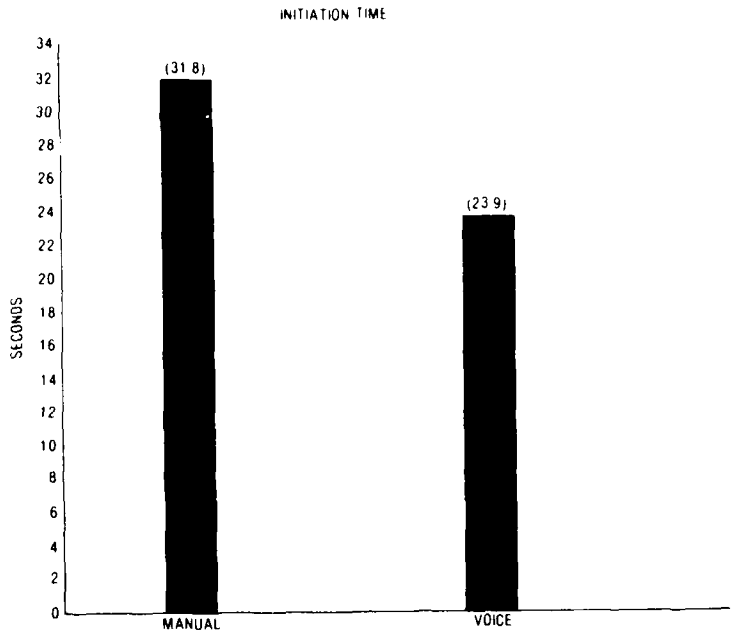


Figure 3

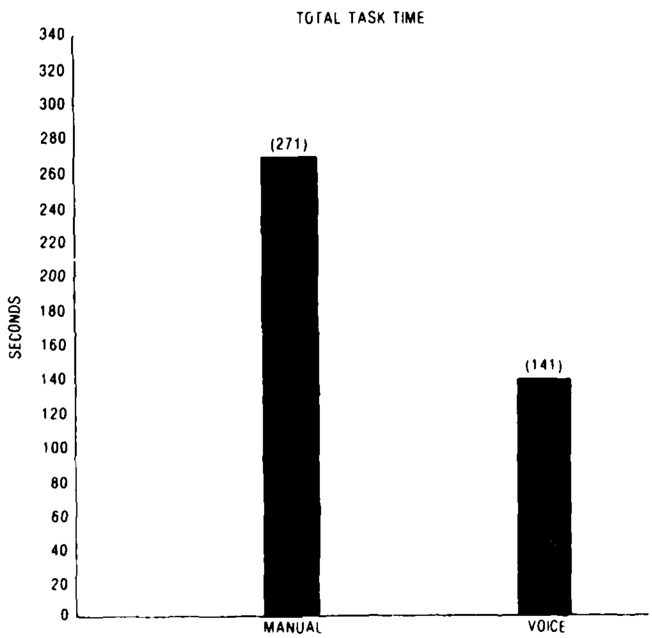


Figure 4

AD-P003 321

The Limits of Multiple Resource Theory in Display Formatting: Effects of task integration

Christopher D. Wickens, University of Illinois; David Boles, Rensselaer Polytechnic Institute; Pamela Tsang, NASA Ames Research Center; and Melody Carswell, University of Illinois

Abstract

Multiple Resource Theory proposes that multi-variable displays will be better processed when information is presented in several different display formats (i.e., verbal, spatial, visual, auditory). We argue here that this superiority of separate formats does not hold if the displayed elements are correlated or must be integrated into a single mental model. We report two experiments that confirm this hypothesis: In experiment 1, subjects monitor two numerical displays. If the Task calls for separate independent decisions concerning the value of each display, performance is best when one is spatial (analog) and the other verbal. If the task calls for integration of both displays into a common decision rule, performance is best when homogeneous display formats are used. In experiment 2 subjects monitor the display of several dynamic elements that compose a dynamic system. Performance is found to be better if these correlated elements are integrated into a configural "object" display, than if they are displayed as separate bargraphs.

Introduction

The concept of multiple resource theory has been proposed as a guideline to formatting displays for complex multi-task systems (Wickens, Sandry & Vidulich, 1983; Wickens, 1984ab). According to this theory the human operator possesses separate processing resources or capacities related to auditory and visual input, and to the processing of spatial and verbal material. Hence the time-sharing efficiency of several tasks will be better if the information for the tasks is displayed using these different formats, than if all information is concentrated within one format. For example, efficiency will improve if the information for one task is displayed visually while that for the other is auditory. In a more general sense the guidelines of multiple resource theory predict that time-sharing efficiency will be greater the farther apart" the two tasks are displayed in some functional space defining human processing capacities.

While the multiple resource guidelines have been well confirmed for displaying information for two separate and independent tasks, the question addressed in the present paper is whether these guidelines also hold when information from two tasks must be integrated into a single "mental model." As an example of different sources of information that must be integrated consider the X, Y & Z positions of several aircraft in an Air Traffic Control

display. It is imperative that the Air Traffic Controller maintain a single mental model of the relative position of all aircraft in the airspace in order to control effectively. In this case it is unlikely that best performance would be obtained when information is distributed across the multiple resource space. Instead, performance should be best when all information is presented via the visual-spatial format (Wickens & Boles, 1983). We report here two experiments that investigate this issue in greater detail.

Experiment 1: Integration of Numerical Information

Method. The purpose of experiment 1 was to determine if the requirement to integrate two numerical indicators, rather than to process them separately, increased the advantage of employing common display codes (i.e., both verbal or both spatial). Twenty-four subjects were presented a series of number pairs. The pair could be displayed in either of three formats shown across the top row of figure 1. In the Heterogeneous format, shown above the left column, one numerical value was indicated by the vertically printed digit name, and the other by a bar graph (either could range in value from 1 to 10). In the two homogeneous formats to the right, the numerical information was presented either as a pair of spatial bar graphs or a pair of verbal digits. Three groups of eight subjects were randomly assigned to one of three tasks to be performed on the digit pair. These tasks are listed down the left edge of figure 1. At the top, the BOTH task was essentially a dual task pair. Subjects had to judge if one number was greater than 5 and if the other was an odd number, and indicate the response to each task with a separate button press. The nature of the two responses are shown in the cells of figure 1.

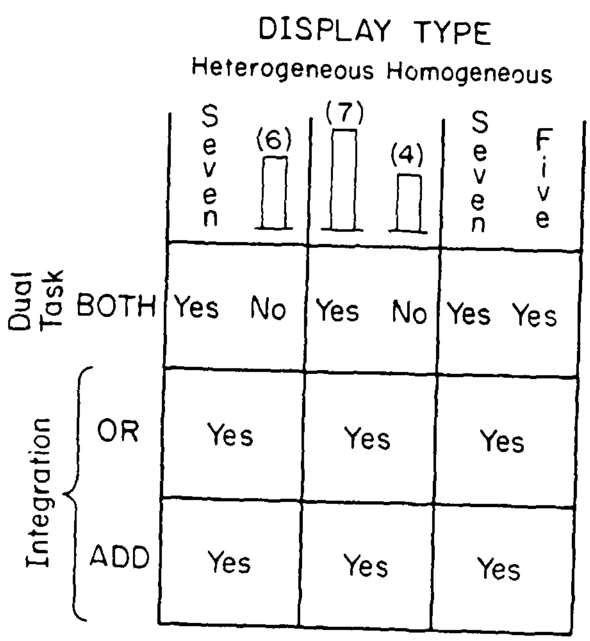


Figure 1: The 3 tasks and 3 display types used in Experiment 1. The nature of the response(s) is shown within the cells of the matrix. The digit over each bargraph representing its value was not present in the actual display

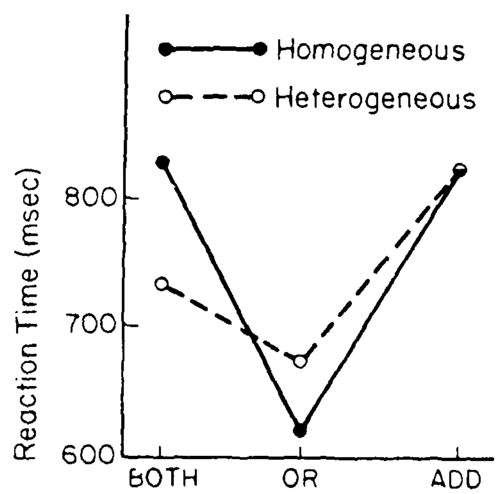


Figure 2: Results of Experiment 1 (averaged over the 2 homogeneous display types).

The bottom two tasks required integration. The OR task required the subject to judge if either digit was greater than 5. The ADD task required the subject to sum the two digits and indicate with a single response if the sum was greater than 10. According to the predictions of multiple resource theory, performance on all three task configurations should be best with the Heterogeneous display configuration, since this configuration distributes the load across verbal and spatial resources. On the other hand the modification of the theory proposed here predicts that this heterogeneous advantage should hold for the BOTH task, but not for the integration ADD and OR tasks.

Results and Discussion. The latency to respond to the stimuli as a joint function of display format and task are shown in figure 2. The data for the two homogeneous display combinations have been averaged, and for the BOTH task only the latency for the first of the two responses is indicated. Figure 2 indicates the predicted interaction of display homogeniety and task type ($F_{2,18}$ = 9.34, p = .002). Decomposition revealed the "crossover" nature of this interaction. In the BOTH task the heterogeneous display generated faster performance as predicted by multiple resource theory ($p < .01$). In the ADD task this advantage evaporated, and in the OR task the homogeneous display combination produced reliably faster performance ($p < .01$). In summary the limitations of the multiple resource theory were confirmed in the predicted direction. The use of separate resources facilitated performance in a dual task situation, but had either no effect, or actually hindered performance when two information sources needed to be integrated or compared before a single response was made.

Experiment 2: Information Integration and the Object Display

Method. In experiment 1, "display proximity" was defined by use of the same display format for both information sources. Experiment 2 investigates the effect of proximity within a display format. All information is presented in analog visual format. However, we contrast a format in which information is displayed on a series of separate bar graphs (low proximity) with a format in which the display dimensions are integrated into a single polygon object, not unlike the iconic display used to present safety parameter information in nuclear power plants (Wood, Wise & Haines, 1981). The task confronted by the subject was to monitor two dynamic systems, each of which contained an output driven by two inputs. The systems were either additive ($O = aI_1 + bI_2$) or

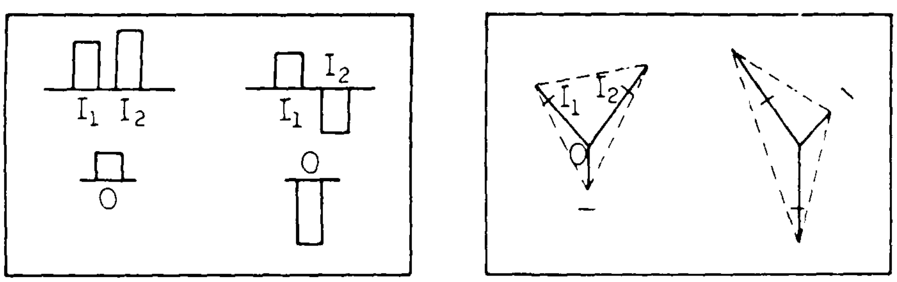


Figure 3: The two systems shown on the bargraph display (left) and on the polygon display (right). The short lines crossing each variable in the polygon display indicate the zero point.

multiplicative ($O = aI_1 \times I_2 + b$). Inputs were slow, semi-random functions uncorrelated with each other. The pair of systems presented on a given display were always both of the same type (multiplicative or additive), but each had different coefficients. Examples of the system pairs presented in each display format are shown in figure 3. At unpredictable intervals, one system or the other would fail. A failure occurred when one of the coefficients would gradually change its value, either increasing or decreasing to zero. The subject's task was to detect which system had failed.

For each system, the task therefore called upon the subject to integrate three sources of displayed information in order to compare these values against a single internal model of the dynamics of a correctly functioning system. As a consequence of this need for integration it was predicted that performance would be superior with the "closer" object display. Eight students at the University of Illinois received extensive briefing about the systems they were to monitor. They were shown the defining system equations, graphic representations of the relation between system variables, and examples of the systems in operation with the occurrence of failures indicated by supplementary visual signals. Following this training, a series of failure detection trials were performed across three sessions.

Results and Discussion. Because of the intrinsic nature of the ramp failures in which the dynamics change became progressively more salient as time went on, practically all failures were eventually detected. Hence the primary dependent variable of interest was latency. These latencies are shown in table 1. It is evident from these data that, as predicted, latency for this integrative type of task is reduced for the "holistic" polygon display relative to the separated bar graph display ($p < .05$). Latencies for the multiplicative systems were shorter than for the additive systems. The advantage of the object display was statistically the same for both system types.

	Additive	Multiplicative
Bargraph	5.1	4.4
Polygon	3.9	3.0

Table 1: Detection Latencies in Experiment 2 (seconds)

In interpreting these results, it is important to realize that the visual angle subtended by the two display formats shown in figure 3 was identical, as were the amplitude and velocity of the variable changes. Hence the advantage to the object display cannot be attributed to the fact that it required less visual scanning or gave rise to more salient display motion. Instead the advantage seems to reflect the greater configural characteristics that resulted when the variables were tied together by the object-defining contours. In terms of the general theory proposed at the outset, the fact that the external inputs were causally related by the system dynamics, and the fact that these variables needed to be integrated in a single mental model, to determine if the systems were functioning normally, produced superior performance when the elements were tied more closely together in visual space.

General Conclusions

Both experiments reported here obtained better performance in tasks requiring information integration when greater proximity between display elements was achieved. In experiment 1 proximity was directly defined in terms of shared processing codes of the multiple resource model. In experiment 2 it was defined in terms of the interconnectedness of display elements. These results in no way invalidate the multiple resource model, but only define limits to its applicability. For the dual task environment, as in the BOTH task of experiment 1, the model is still perfectly valid. Furthermore, the boundary conditions that define the limits can be specified a priori. Task integration defines formally a situation in which the optimal response to one stimulus or display variable cannot be specified without specifying the level of another variable. Models that prescribe optimal system design are rarely applicable under all circumstances. Once such models are developed it is necessary to specify their boundary conditions. The present research has attempted to perform this function.

Acknowledgements: This research was supported by a contract from the Office of Naval Research Engineering Psychology Program, N-000-14-79-C-0658. Gerald Malecki was the technical monitor.

References

Wickens, C.D. Processing Resources in Attention. In R. Parasuraman & R. Davies (Eds.) Varieties of Attention. San Diego: Academic Press, 1984(a).

Wickens, C.D. Engineering Psychology & Human Performance. Columbus, Ohio: Charles Merrill, 1984(b).

Wickens, C.D. & Boles, D. The Limits of Multiple Resource Theory. University of Ill. Engineering Psych. Lab Tech Report EPL-83-5/ONR-83-4. November 1983.

Wickens, C.D., Sandry, D., & Vidulich, M. Compatibility & Resource Competition. Human Factors, 1983, 25, 227-248.

Wood, D., Wise, S., & Hanes, L. An evaluation of Nuclear Power Plant Safety Parameter Display Systems. In R. Sugarman (Ed.) Proceedings 25th Annual Meeting Human Factors Society. Santa Monica: Human Factors, 1981.

AD-P003 322

TRACKING WITH INTERMITTENT RADAR COVERAGE:

I. INTERRUPTIONS AFTER EACH COLLECTED FRAME OF IMAGERY*

John R. Bloomfield and Raleigh K. Little

Honeywell Systems and Research Center
2600 Ridgway Parkway
Minneapolis, Minnesota 55413

Abstract

Imagery simulating the output of a moving target indicator (MTI) radar was used in two experiments to determine the effect on the ability of radar operators to track targets moving in one particular area, when the radar was being switched intermittently to provide coverage of a second area as well. In both experiments, we investigated the effects of switching away from the area of interest after each complete scan of that area, varying the length of each interruption that occurred before returning for another scan of the primary area. The frames of simulated MTI radar imagery used for each condition were shown in time compression to twelve operators in both experiments. The results of the two experiments seem to present a coherent picture. Increasing the length of the interruptions in coverage from 15 to 30 seconds had little effect on tracking performance. With interruptions of 45 to 90 seconds, performance was worse with smaller target units, containing ten vehicles, than it was for larger units. However, with interruptions of 120 seconds, there was a decrement for all target units, large and small, compared to the level of tracking performance achieved with shorter interruptions.

Introduction

In a previous study carried out as part of the Stand-Off Target Acquisition System (SOTAS) program, Bloomfield (1981) investigated the effects on detection performance of varyfng various display parameters. The imagery presented to the operators who took part in that study simulated the output of a moving target indicator (MTI) radar. The radar was assumed to be scanning the particular area of interest once every fifteen seconds, so that the data presented to the operators was updated four times a minute.

In a series of studies, the first of which is reported here, we explored the situation in which a similar MTI radar would be used to provide coverage of more than one area of interest. If the radar were to be switched after each individual scan, there would be an update on both areas of interest every thirty seconds, although the two updates would be phased fifteen seconds apart. If the radar were to be switched after two or more scans of either area, the update

*This work was carried out for the U.S. Army Electronics and Research Development Command under contract NO. DAAK20-80-C0254.

schedule would be more complex. In the two current experiments, we investigated the effects of switching away from the area of interest after each complete scan of that area. We varied the length of the interruptions that occurred before returning for another single scan of the primary area.

In both experiments, the imagery was presented in time compression. With this mode of presentation, first demonstrated by White (1956), several previously-obtained frames of radar imagery are collected and stored. Then, they are played back to the operator in the same order in which they were acquired, but at a faster rate. Returns from a particular moving target appear on the screen as a steadily moving dot.

Method

Imagery: The MTI radar imagery used in this study was devleoped and shown on the ground station simulator designed by Lorence, Kittleson, Ottoson, Baum, and Graffunder (1980). Basically this simulator consists of the hardware and software necessary for simulating a radar ground station, and of a set of computer models capable of simulating radar returns and of converting those returns into imagery.

The imagery used in this series of experiments was taken from a set of 27 specially produced master sequences of simulated MTI radar frames. Each master sequence consisted of 96 frames of imagery assumed to have been collected at fifteen-second intervals throughout a 24-minute period of time.

We wanted our experimental imagery to be representative of the kind of imagery that would be collected by an MTI radar in a realistic operational setting. At the same time, we needed to control the characteristics of the target and some nearby units. These goals were not completely compatible. However, by basing our new imagery on the Stand-Off Target Acquisition System Movement Scenario developed by Stephens, Dolby, Plocher and Little (1980), we were able to satisfy each of them to a large extent. This Scenario was derived from the U.S. Army's depiction of a potential attack on western Europe by Warsaw Pact force (see SCORES, Europe III). It provides a detailed account of expected enemy tactical movement over an eleven-hour period.

Three target areas were chosen from the Movement Scenario battlefield. Within each area, a highly-travelled route was identified. We determined which segments of those routes were masked from the radar, so we could select target paths that varied in the extent to which they were masked on all three routes.

A tracking task was used in the experiments. On each trial the operator was asked to track a designated target as it moved along its particular path. This was not easy, both because of the masking, and because of the presence of other simulated units moving in the vicinity. On each trial, there were four of these other, distractor units. They were of the same size, containing 10, 25, or 50 vehicles, were in the same single-column formations, and moved at the same speed, 10, 20, or 30 kph, as the target unit. In addition, the paths travelled by these units were selected in such a way that they would either intersect, or come close to, the path of the target unit at a point in the trial where it was passing through a masked area. Deliberately, these units were made difficult to distinguish from the target. Fewer cues were provided than might be expected in most operational situations. The operator had to rely on the coherence of motion, and direction of that motion, in deciding which of the units emerging from a masked area were the distracting units and which the designated target.

Design: The basic experimental design was a four-way analysis of variance factorial design. For both experiments, four interruption conditions were combined with the three target areas, three target unit sizes, and three target unit velocities to give 108 unique trials. All 108 trials were presented to each operator taking part in each experiment. The trials were presented in a different, random order to each operator.

The interruption conditions investigated in both experiments were obtained by omitting various frames from the 27 master sequences of imagery. These conditions are shown in Tables 1 and 2.

INTERRUPTION CONDITION	RATE OF UPDATE (SECONDS)	NUMBER OF FRAMES OMITTED	FRAME NUMBER																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	15	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
B	30	1	X		X		X		X		X		X		X		X		X
C	60	3	X				X				X				X				X
D	120	7	X								X								X

Table 1. First Experiment: Interruptions of 0, 1, 3 or 7 Frames

INTERRUPTION CONDITION	RATE OF UPDATE (SECONDS)	NUMBER OF FRAMES OMITTED	FRAME NUMBER																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
E	45	2	X			X			X			X			X			X	
F*	60	3	X				X				X				X				X
G	75	4	X					X					X					X	
H	90	5	X						X						X				

*Note: Condition F in this experiment is identical to Condition C in the first experiment

Table 2. Second Experiment: Interruptions of 2, 3, 4, or 5 Frames

The tables show the update rate for each condition, the number of consecutive frames omitted, and then the particular frames retained of the first 17 in the 96-frame master sequences.

Operators: The operators used in these experiments were selected from a pool of 21 veterans (10 ex-Army, 6 ex-Air Force, 3 ex-Navy, and 2 ex-Marine). They were selected because they closely match the population of Army radar operators. Twelve operators took part in the first experiment. For the second experiment, nine of the same twelve operators were used again, with three new people.

Procedure: At the start of each experimental session, the operator was seated in front of a video display terminal, while the experimenter read the instructions for the experiment and answered any questions the operator might have. At the beginning of each trial the initial frame of imagery was displayed and a cursor was centered on the lead vehicle of the target. The imagery was presented in time compression, at a rate of five frames per second. At the end of the trial, the last frame in the trial sequence remained on the screen, and the operator indicated where he or she thought the target unit had stopped by moving the cursor. Trials could not be repeated. The operator was not told whether the responses were correct or incorrect.

Results and Discussion

In both experiments, for each operator, there were nine trials with each combination of interruption condition and target unit size. These trials were the product of using three target areas with three target unit velocities. The results were averaged across the twelve operators taking part in each experiment.

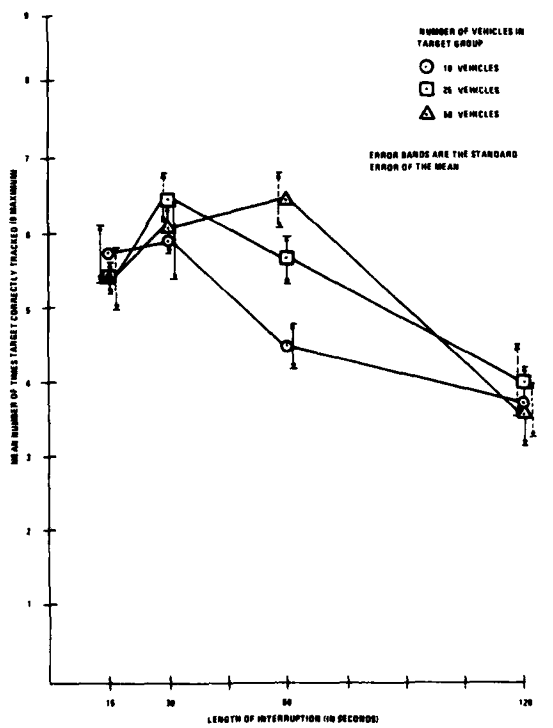


Figure 1. Tracking Performance as a Function of Duration of Interruption: Data Averaged Across Twelve Operators.

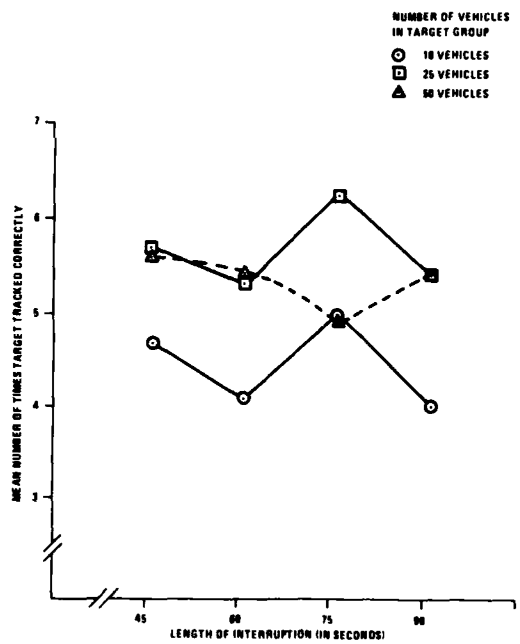


Figure 2. Tracking Performance as a Function of Interruption: Data Averaged Across Twelve Observers.

Figure 1 shows a plot of the mean number of successfully tracked targets as a function of the duration of the interruptions in the first experiment. When the duration changed from 15 to 30 seconds, there was little effect on tracking performance. With a 60-second interruption, performance deteriorated for the smaller, 10-vehicle target units, but not for the larger units, producing a statistically significant interaction between unit size and interruption duration at the $p < .05$ level. When the interruptions were as long as 120 seconds, there was a decrement for all target units, irrespective of size: the effect of interruption duration was significant at the $p < .0005$ level.

Figure 2 shows a similar plot for the second experiment. The interruption conditions investigated in this case did not significantly effect tracking performance, although the effect of target unit size was significant at the $p < .0005$ level.

The results of the two experiments seem to present a coherent picture. The first experiment showed there was little effect on tracking performance when the duration of the interruptions changed from 15 to 30 seconds. Then, for interruptions of 45 to 90 seconds, performance was worse with smaller target units, with 10 vehicles, than it was for the larger units (second experiment, confirmed in first experiment for the 60-second interruption condition). Finally, with an interruption of 120 seconds, there was no difference in tracking performance that was attributable to unit size, but, there was a decrement for all targets when compared with the performance level achieved with the shorter interruptions (first experiment).

REFERENCES

Bloomfield, J. R. Trade-Off Studies: SOTAS Imagery. Honeywell Systems and Research Center, Minneapolis, Minnesota. Technical Report No B010, prepared for U.S. Army Electronics Research and Development Command, under Contract No. DAAK20-79-0278. 1981 (CONFIDENTIAL)

Lorence, L. R., Kittleson, C. G., Ottoson, K.O., Baum, D. R. and Graffunder, K. SOTAS Ground Station Simulator Design Plan. Honeywell Systems and Research Center, Minneapolis, Minnesota. Technical Report prepared for U.S. Army Electronics Research and Development Command, under Contract NO. DAAK20-80-C-0254. 1980.

Scores Europe III (CACDA Manual Wargaming Report, Standard Scenario for Combat Developments. Europe III, Sequence 2A. ACN 24826. Ft. Leavenworth, Kansas. 1980.) (SECRET)

Stephens, J. A., Dolby, M. B., Plocher, T.A., and Little, R. K. Functional and Operational Characteristics of the Stand-Off Target Acquisition System (SOTAS): Simulation Update/SOTAS Movement Scenario. Honeywell Systems and Research Center, Minneapolis, Minnesota. Technical Report prepared for U.S. Army Electronics Research and Development Command, under Contract No. DAAK20-79-C-0278. 1980. (CONFIDENTIAL)

White, C. T. Time Compression of Radar and Sonar Displays. Naval Electronics Laboratory Report No. 671. San Diego, California. 1956.

AD-P003 323

TRACKING WITH INTERMITTENT RADAR COVERAGE

II: INTERRUPTIONS AFTER TWO OR MORE CONSECUTIVELY-COLLECTED FRAMES OF IMAGERY*

Raleigh K. Little and John R. Bloomfield

Honeywell Systems and Research Center
2600 Ridgway Parkway
Minneapolis, Minnesota, 55413

Abstract

Simulated moving target indicator (MTI) radar imagery was used in the two experiments performed in this, one of an on-going series of studies in which we are investigating conditions of intermittent radar coverage. In the first of these experiments there were two, and in the second experiment there were two, three, or four consecutively-collected frames of imagery before each interruption in coverage. In both experiments there were twelve operators and the imagery was presented in time compression. The results of the two experiments showed that the use of more than one frame of imagery before the occurrence of the interruptions in radar coverage resulted in improved performance. For the larger, 25- and 50- vehicle units, two consecutively-collected frames, and for the smaller 10- vehicle units, four consecutivelycollected frames helped to prevent the decrement in tracking performance that otherwise occurred with 120-second interruptions.

Introduction

This is the second study in a series of investigations of the situation in which a moving target indicator (MTI) radar is switched back-and-forth between two areas, in order to provide coverage of both of them. In the first study, Bloomfield and Little (1984) investigated the effects of switching away from one area after each complete radar scan of that area and varying the length of the interruptions. They found that there were decrements in the tracking performance of operators with small, 10- vehicle target units with interruptions of 45 seconds or longer, and with larger, 25-, or 50- vehicle units with interruptions of 120 seconds. The two experiments reported here used imagery from the master sequences developed by Bloomfield and Little, and were performed in order to determine whether increasing the number of consecutively-collected frames before each interruption would reduce or eliminate those decrements.

It should be noted that, the fact that there were interruptions in radar coverage, did not mean that there were corresponding interruptions in the flow of images being displayed to the operator. He or she did not see any blanks or gaps. After selected frames were removed from the master sequences, in order to produce the various experimental conditions, the remaining frames were shown in time compression. With this mode of presentation, several frames of imagery are collected then played back in the same order that they were acquired, but at a faster rate. The result is that returns from a particular moving target appear on an operator's display as a steadily moving dot. When only single frames were retained before the interruptions, as was the case in both

* This work was carried out for the U.S. Army Electronics and Research Development Command under Contract No. DAAK20-80-C-0254.

experiments in Bloomfield and Little's study, the effect was to lengthen the update rate, reduce the number of images to be viewed and shorten the viewing time. In the two experiments reported here, another effect was possible. The use of or more consecutive frames before an interruption results in a variable update rate, as can be seen in Table 1, which shows the interruption conditions for the first experiment. It was possible that, when presented in time compression, imagery obtained with variable update rate would produce target movement that appeared to be spasmodic rather than smooth, and that target units might be harder to track.

FIRST EXPERIMENT: INTERRUPTIONS AFTER TWO CONSECUTIVELY-COLLECTED FRAMES

Method. The generation of the master sequences of simulated MTI imagery was described by Bloomfield and Little. We obtained our experimental conditions by omitting selected frames from these master sequences. The conditions are shown in Table 1.

INTERRUPTION CONDITION	RATE OF UPDATE (SECONDS)	NUMBER OF FRAMES OMITTED	FRAME NUMBER																			
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	15	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
B	15,30	1	X	X		X	X		X	X		X	X		X	X		X	X		X	X
C	15,60	3	X	X				X	X				X	X				X	X			
D	15,120	7	X	X								X	X								X	X

TABLE 1. Interruptions of 0, 1, 3, or 7 Frames Occurring After Each Pair of Consecutively-Collected Frames

The conditions used here were similar to those used in Bloomfield and Little's first experiment: the interruptions were of the same duration, but two frames were retained before each interruption, instead of one.

The imagery was shown in time compression, with a presentation rate of five frames per second. The experimental procedure was essentially the same as that described by Bloomfield and Little.

Twelve operators were used. Eight of them took part in both of Bloomfield and Little's experiments, three took part in their second experiment, and one was new.

Results. For each of the twelve operators, there were nine experimental trials for each combination of four interruption conditions and three target unit sizes. Figure 1 shows a plot of the average number of trials on which targets were successfully tracked, as a function of these combinations.

An analysis of variance performed on this data showed the interaction between the two main effects, interruption duration and target size, was significant at the $p < .0005$ level. The interaction can be seen in Figure 1. For the larger

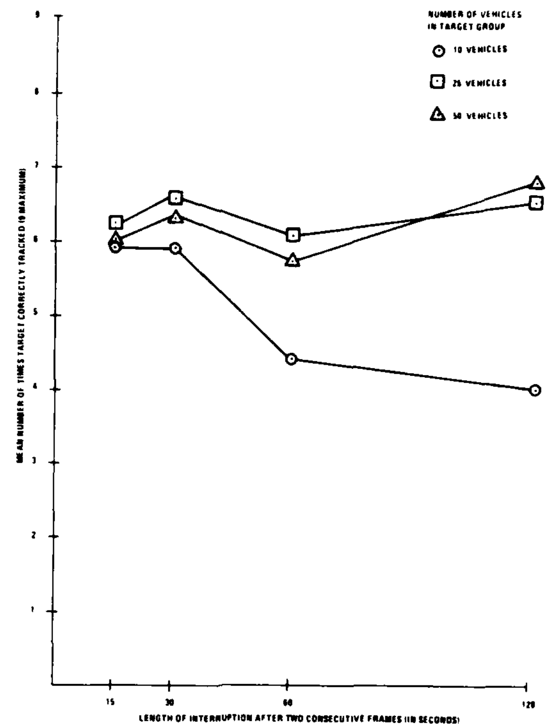


Figure 1. Tracking Performance as a Function of Duration of Interruption After Two Consecutive Frames: Data Averaged Across Twelve Operators.

units, with 25 or 50 vehicles, changes in the duration of the interruptions up to 120 seconds had no effect. On the other hand, for smaller, 10-vehicle units, a decrement occurred with 60-and 120-second interruptions.

The fear that variable update rates might lead to displayed targets appearing to move in a spasmodic fashion and, as a result, being harder to track, proved to be unfounded, with the particular update rates used in this experiment.

When this experiment is compared with Bloomfield and Little's first experiment, some interesting differences and similarities emerge. First, the differences: in the earlier experiment, there were decrements in performance for the two larger target units for 120-second interruptions, while here, with two consecutive frames retained before each interruption in coverage, those decrements have disappeared. Next, the similarities: with the smaller, 10-vehicle target units, tracking performance was significantly poorer with 60- and 120-second interruptions, both when there was one, and when there were two frames of imagery before each interruption. The addition of the second frame did not help. However, it seemed possible that the addition of further frames, so there would be three or four of them before each interruption, might be beneficial. The next experiment explored this possibility.

SECOND EXPERIMENT: INTERRUPTIONS AFTER THREE OR FOUR CONSECUTIVELY-COLLECTED FRAMES

Method: Several changes in experimental design were made for this experiment. Two target unit sizes, the larger, with 50 vehicles, and the smaller, with 10, were used instead of three: the 25-vehicle units, which had produced tracking performance very like that obtained with the 50-vehicle units, were omitted. In addition, only two interruption durations were used: 30 and 120 seconds. There were, however, six interruption conditions. They are shown in Table 2

INTERRUPTION CONDITION	RATE OF UPDATE (SECONDS)	RATIO OF COLLECTED TO OMITTED FRAMES	FRAME NUMBER																									
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
A	30	1/1	X		X		X		X		X		X		X		X		X		X		X		X		X	
B	15,30	2/1	X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X
C	120	1/7	X								X								X								X	
D	15,120	2/7	X	X								X	X								X	X						
E	15,15,120	3/7	X	X	X								X	X	X							X	X	X				
F	15,15,15,120	4/7	X	X	X	X								X	X	X	X							X	X	X	X	

Table 2: Interruptions of 1 or 7 Frames After 1, 2, 3, or 4 Consecutively - Collected Frames

It should be noted that conditions A and C were similar to two conditions (B and D) in Bloomfield and Little's first experiment. And conditions B and D were similar to conditions B and D in the first experiment reported here. Conditions E and F, with three and four consecutively-collected frames, had not been tested before.

The imagery was presented in time compression to twelve operators, ten of whom had taken part in earlier experiments in this series, and two of whom were new.

Results. While the duration of the interruptions and the number of consecutive frames were both varied, a full factorial experiment was not conducted. Hence, an overall analysis of the data was not justified. Instead the results were examined by means of several detailed, partial comaprisons. A series of t-tests were performed using Dunn's procedure to control error rate (Kirk, 1968).

Two comparisons of the effect of changing the duration of interruption were made. These involved comparing performance with 30- and 120- second interruptions and one and two consecutive frames. With one frame, the reductions in tracking performance from 30 (condition A) to 120 seconds (condition C) were not large enough to be statistically significant, as they were in Bloomfield and Little's first experiment. It is possible that the higher-than-expected scores obtained here for these two 120-second interruption conditions may, in part, have been the result of practice. Ten of the twelve operators had taken part in one, or more, previous experiments in this series. Also, since the same basic imagery was used throughout the series, and, in addition, the particular characteristics of conditions A and C had been utilized before, it is possible that some learning

could have occurred. Some support for this possibility is suggested by the fact that the two operators who had not taken part in any of the earlier experiments in the series did perform at a level comparable to Bloomfield and Little's operators.

With two consecutive frames, with the larger units there was no decrement in performance, while with the smaller, significantly fewer (at the $p < .01$ level) target units were tracked successfully with the 120-second interruption, confirming the results of the first experiment reported here.

Two comparisons of the effect of increasing the number of consecutively-collected frames of imagery were made: one with 30-second, the other with 120-second interruptions. As might be expected from a comparison of the first experiments in this report and in Bloomfield and Little's study, tracking performance is similar, whether there there are one or two consecutive frames, for both large and small target units, when there are 30-second interruptions in radar coverage.

With 120-second interruptions, the main point of interest was, whether the use of three or four consecutive frames of imagery before the interruptions, would lead to improvements with the smaller target units. When there were two or three consecutive frames, the mean number of 10- vehicle units successfully tracked was significantly smaller than the number of 50- vehicle units (at the $p < .05$ level). However, with four consecutive frames, the difference between the number of small and large target units successfully tracked was not significant. The use of four consecutive frames did result in improvements.

CONCLUSION

The presentation of two or more consecutive frames of imagery, instead of only one, before the occurrence of the interruptions in MTI radar coverage, resulted in improved tracking performance. With larger target units, containing 25 or 50 vehicles, there was no decrement in performance, when two consecutive frames of imagery were used with interruptions as long as 120 seconds (first experiment). For the smaller units, with 10 vehicles, the use of four consecutive frames of imagery reduced the decrement in performance, that otherwise occurred with a 120-second interruption (second experiment).

So far, this series of experiments has shown that, as far as tracking performance is concerned, there is no particular advantage in updating the MTI imagery more frequently than every 30 seconds. Further, if four consecutive frames can be acquired before the radar is switched to cover another area, interruptions of up to 120-seconds can be sustained without a statistically significant decrement in performance.

REFERENCES

Bloomfield, J.R. and Little, R.K. Tracking with Intermittent Radar Coverage I. Interruptions After Each Collected Frame of Imagery. Previous Paper in these Proceedings. 1984.

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Brooks-Cole/Wadsworth. Belmont, California. 1968.

AD-P003 324

Family and Work in the Air Force

Mickey R. Dansby
Air Command and Staff College

Janice M. Hightower
Leadership and Management Development Center

Abstract

Multiple regression models relating attitudes on job performance, job related satisfaction, and retention in the Air Force to family variables are presented. Variance accounted for ranged from 12% To 55%, with member's perceptions of family attitudes and spousal attitude toward retention contributing heavily.

In recent years, family issues have become increasingly important to the Air Force. This interest reflects many concerns, among which are the impact of the all volunteer force, problems in recruiting and retaining personnel, and the effects of demographic changes, reflective of society as a whole, in Air Force families. Several major research efforts, such as the Families in Blue studies (Orthner, 1980; Orthner & Bowen, 1982), and comprehensive literature reviews (Black, 1982; Hunter, 1982) have highlighted some critical research concerns in Air Force family matters.

Despite the increased concern for family issues, many questions remain unanswered. Recently, the Air Force Leadership and Management Development Center (LMDC) at Maxwell AFB, Alabama, developed the U.S. Air Force Spouse Survey (AFSS) as a research instrument to help investigate the relationships between family attitudes and job and retention variables (Ibsen & Austin, 1983). An important feature of the survey is that it matches spousal responses with the responses of the military member on another survey, the Organizational Assessment Package (OAP), which measures the member's attitudes on a number of relevant job and retention dimensions. Thus, one is able to examine direct correlations between the attitudes of military members and their spouses. Based on OAP and AFSS results, the present research develops models relating attitudes on job performance, job related satisfaction, and retention to family issues.

Variables Included in the Models

Many factors have been proposed as family influences on the military member's job performance, satisfaction, and decision to remain in the military service. Not surprisingly, the factor suggested most often is the influence of the spouse's attitudes toward, and commitment to, the military lifestyle (Lund, 1978; Orthner, 1980; Szoc, 1982). Spousal attitudes are particularly important in influencing the retention decision (Lund, 1978; Orthner, 1980). Spousal support also appears to influence job performance and satisfaction (Schneider & Dachler, 1978). Other family variables suggested as influences on retention include

family separations (Lund, 1978), member's perceived impact of work on the family (Schneider & Dachler, 1978), dual career considerations (Black, 1982), traditional versus egalitarian family models (Black, 1982), family integration into the military community (Wood, 1982), time for family interaction (Szoc, 1982), quality of life and family income (Szoc, 1982), and family disruptions caused by the job, such as frequency of moves (Szoc, 1982).

In addition to the factors mentioned above, there are several general family variables which researchers and theorists suggest may affect job variables. Included among these are family life cycle, stress, and cohesion (McCubbin, Cauble, & Patterson, 1982); family size, coping style, and commitment to the Air Force (Orthner & Bowen, 1983); family structure variables such as single parenthood, etc. (Orthner, 1980); feelings of status, spousal autonomy, and family patterns of external social interaction (Schneider & Dachler, 1978); marital satisfaction (Szoc, 1982); and a variety of family economic and support (e.g., medical care) concerns (Schneider & Dachler, 1978; Szoc, 1982).

Based upon these studies, a general model was constructed for predicting job related satisfaction (JRS), perceived work group effectiveness (PWGE), and career intention (CI) to remain in the Air Force (all measured by the OAP) from family variables (as measured by the AFSS and OAP). After reduction of variables in the AFSS via factor analysis, a multiple regression model was developed including the following predictor variables from the AFSS: time on present station; length of marriage; number of children; number of children at home; spousal employment patterns; compatibility of spouse's and member's work schedules; spousal identification with the Air Force; perceived economic security; family separation frequency and duration (i.e., temporary duty [TDY] assignments); perceived influence of temporary separations on family life; satisfaction with job benefits; satisfaction with recreation services; spousal identification with the military job; satisfaction with basic family services; perceived time pressure from the job; spouse's interest in the member's job; satisfaction with medical/dental services; perceived job-provided status; attitude toward move frequency; spousal desire for career retention; congruence between spouse's and member's desire for a military career; perceived demands on spouse to "participate" in activities enhancing member's career; perceived importance of spousal attitudes toward member's career decisions; and perceived equity of compensation. One predictor variable from the OAP, member's perception of family attitude toward the job, was also included.

Results

The regression models were developed based on the matched responses of 4337 Air Force workers and their spouses. The workers included: 21% officers, 67% enlisted, and 12% civil servants; 20% with less than four years of service, and 37% between 4 and 12 years; 23% with less than one year on station, and 48% between one and three years; 48% with less than one year in their present jobs; 12% minority group members; 3% unmarried, 11%

married to military spouses, 34% married to civilian employed spouses, 52% married to unemployed spouses, and 3% separated geographically from their spouses; 29% college graduates; 50% supervisors of at least one person; 32% on shifts other than the normal day shift; 15% rated personnel; and 90% males. The spouses included: 25% age 25 or under, 50% age 30 or under, and 90% age 42 or under; 30% married less than four years, 49% less than eight years, and 91% less than 20 years; 52% living off base; 12% minority group members; 17% college graduates; 25% with no children at home, 24% with only one child at home, and 35% with two children at home; 72% (of those who worked) working normal day shift, and 40% of working spouses working due to financial necessity.

Regressions for the full models obtained the following results: (1) for JRS, $\underline{R}$ =.74 ($\underline{R\text{-square}}$ =.55), $\underline{p}$ <.001; (2) for PWGE, $\underline{R}$ =.36 ($\underline{R\text{-square}}$ =.13), $\underline{p}$ <.001; and (3) for CI, $\underline{R}$ =.53 ($\underline{R\text{-square}}$ =.28), $\underline{p}$ <.001.

Restricted models were constructed using stepwise inclusion where the criterion for inclusion was $\underline{p}$ <.05. Regressions were as follows: (1) for JRS, $\underline{R}$ =.74 ($\underline{R\text{-square}}$ =.55), $\underline{p}$ <.001; (2) for PWGE, $\underline{R}$ =.35 ($\underline{R\text{-square}}$ =.12), $\underline{p}$ <.001; and (3) for CI, $\underline{R}$ =.53 ($\underline{R\text{-square}}$ =.28), $\underline{p}$ <.001. One predictor variable, member's perception of family attitude toward the job, appeared in all three restricted models. Predictors included in the regressions for both JRS and PWGE were number of children, spousal identification with the military job, and spouse's interest in the job. Predictors unique to the JRS regression included: perceived time pressure from the job, compatibility of spouse's and member's work schedules, whether or not the spouse was employed, and attitude toward move frequency. Predictors unique to the PWGE regression included: time on present station, family separation (TDY) duration, and perceived equity of compensation. Finally, predictors unique to CI included: spousal desire for career retenton, length of marriage, number of children at home, congruence between spouse's and member's desire for a military career, perceived influence of family separation on family life, satisfaction with recreation services, family separation (TDY) frequency, perceived economic security, and satisfaction with job benefits.

Beta weights indicated that member's perception of family attitude toward the job was the most important contributor in the restricted regression models for JRS and PWGE. The second heaviest contributor for JRS was spousal identification with the military job. All other variables in the restricted JRS model contributed roughly equally. For PWGE, all variables other than perception of family attitude contributed about equally to the restricted model. The greatest contributor in the restricted model for CI was spousal desire for career retention. This was followed by length of marriage, member's perception of family attitude toward the job, and number of children at home, all with similar weights. The remaining variables in the CI model all contributed lesser, and similar, weights.

Regression models excluding the member's perception of family

attitude toward the job were also derived. These regressions predicted JRS, PWGE, and CI based only on spousal responses. Substantial reductions in predictive ability accrued for JRS ($\underline{R}$ =.42, $\underline{p}$ <.001) and PWGE ($\underline{R}$ =.20, $\underline{p}$ <.001), but not for CI ($\underline{R}$ =.51, $\underline{p}$ <.001). In these regressions, the heaviest contributors were spousal identification with the military job (for JRS and PWGE), and spousal desire for career retention (for CI).

Examination of the simple regressions for predictor variables on JRS, PWGE, and CI showed that predicted scores all improved as the member's perception of his/her family's attitude toward the job improved. Likewise, as spousal identification with the job increased, so did JRS and PWGE scores. There were also positive correlations between number of children and both JRS and PWGE. Time pressure contributed negatively to JRS. Finally, for CI, intention to stay in the Air Force correlated positively with spousal desire for the member to stay in, length of time married, and number of children at home.

Discussion

These results are encouraging for further exploration of the relationships between job and family. Although there are significant relationships between the two, causal sequences are not clear. Does the family's attitude toward the Air Force job cause the member to develop similar attitudes, or is it the member's attitude that infects his/her family? Perhaps the relationship is synergistic, with a mutual evolution and reinforcement of attitudes between the member and the family. Research (e.g., Szoc, 1982) suggests negative family attitudes toward the member's military career, no matter how these attitudes develop, have a negative impact on the member's career intentions. Our assessment is that military leaders would be wise to continue efforts to improve family life in the Air Force as a way of improving productivity and retention.

The present research suggests several steps leaders might take, where possible, to improve family life. For example, leaders might try to increase spousal identification with the job by expanding the interface between the job and the family (e.g., open houses, etc.). Consideration could be given to more time off to be with family and to scheduling members so that member/spouse work schedules are compatible. Family moves and TDY separations should be held [illegible] minimum. Attention should also be directed toward improvi[illegible] [illegible]nomic security, recreation services, and medical/dental [illegible].

Although the [illegible]s presented here show significant relationships bet[illegible] family and work variables, the amount of variance accounted for is small to moderate (55% for JRS, 12% for PWGE, and 28% for CI). Other family variables not assessed in the current AFSS, such as marital satisfaction, may be needed to improve the models. Furthermore, one should not forget that work influences such as supervisory, co-worker, and task variables contribute significantly to JRS, PWGE, and CI. Since removal from

the models of the member's perceptions of family attitudes substantially reduced predictive ability for JRS and PWGE, such work-related variables may be more important in influencing the worker's perception of how his/her family views his/her work than are the actual views of the family. Further research is needed to clarify the relative contributions of work and family variables to the productivity and retention of Air Force personnel.

References

Black, L. W. Changing patterns of Air Force families. (Report 82-0275). Maxwell AFB, Alabama: Air Command and Staff College, March 1982.

Hunter, E. Families under the flag. New York: Praeger, 1982.

Ibsen, K. & Austin, J. Initial development of the U.S. Air Force Spouse Survey. Paper presented at the meeting of the Military Testing Assoc., Gulfport, Mississippi, October 1983.

Lund, D. A. Junior officer retention in the modern volunteer army: Who leaves and who stays. In E. Hunter & D. Nice (Eds.), Military families: Adaptation to change. New York: Praeger, 1978.

McCubbin, H. I., Cauble, A. E., & Patterson, J. M. Family stress, coping, and social support. Springfield, Illinois: Charles C. Thomas, Publisher, 1982.

Orthner, D. K. Families in blue: A study of married and single parent families in the U.S. Air Force. Greeensboro, North Carolina: Family Research and Analysis, Inc., 1980.

Orthner, D. K. & Bowen, G. L. Families in blue: Phase II. Insights from families in the pacific. Washington, D.C.: Human Resources Research and Development Center, SRA Corporation, 1982.

Orthner, D. K. & Bowen, G. L. Families in blue: Opportunities for Ministry. Arlington, Virginia: Human Resources Research and Development Center, SRA Corporation, 1983.

Schneider, B. & Dachler, H. P. Work, family, and career considerations in understanding employee turnover intentions. (Research Report #19). College Park, Maryland: University of Maryland, 1978.

Szoc, R. Family factors critical to the retention of Naval personnel. Columbia, Maryland: Westinghouse Public Applied Systems, 1982.

Wood, F. Work-family conflict of Air Force junior officers. Doctoral dissertation, U.S. Air Force Academy, Colorado, 1982.

AD-P003 325

The Changing Roles of Military Wives: An Analysis of Work-Related Values and Lifestyle Variables

Lorie L. Mazzaroppi
College of Business Administration at Madison
Fairleigh Dickinson University

Abstract

This paper reports the preliminary results of a survey to investigate the significance of the changing roles of military wives who are attempting to find a balance between their desire for the responsiblity, satisfaction, and economic independence of a job/career, and the desire to have children and a home life.

The sample for the analysis consists of 75 civilian military spouses who responded anonymously to a questionnaire, accompanied by a transmittal letter, distributed through the office of the "Military Family Member Employment Program" established under a Civilian Personnel Office-sponsored effort at the Fort Monmouth Army Base in New Jersey. One purpose of the study is to provide the Military Family Member Employment Coordinators with a broad overview of the client users of the Program. This, in turn, will serve to enhance the effectiveness of the Program as it relates to planning and programming employment counseling services geared specifically to meet the needs of military spouses.

Introduction

The transformation of our society from one in which most women stayed home most of the time to one in which most women work has wrought historic alterations not only within the domestic household but within the workplace as well (Dowd, 1983; Kagan, April 1983). Of all the changes in the workplace none has been more significant than the dramatic influx of women into the job market. For example, in March, 1981, about 25.5 million wives, or 51 percent, were working or looking for work (Hayghe, 1982). In 1982, the labor force participation rate was 53 percent with 47.1 million women in the workplace, and is expected to rise to 65 percent by 1995 (Shellenbarger, 1982; U.S.News & World Report, 1982).

The data indicate that working women, working parents, dual career and dual worker families are a growing segment of American society and are the wave of future employment (Hayghe, 1982; Shellenbarger, 1982; Quinlan, 1980). As a consequence of these trends, the family is taking on new forms. The traditional nuclear family now accounts for only 20 percent of households in the United States. Families are in a significant state of flux and the uncertainty reverberates throughout society. The dual-work family and the single parent family are having the most impact on the relationship between the work organization and the family. Both forms of families have increased significantly within the last ten years (Hunsaker, 1983). This study focuses on one segment of this new phenomenon in the American family, that is, the dual-income military family household. In particular, this research concentrates on the civilian spouses of military members.

Studies have shown that military families are also clearly being influenced by shifts in family values today. Attitudes toward parental roles, family leadership, household responsibilities and employment of wives are all undergoing changes. Many men and women are seeking new balances between work and family responsibilities (Orthner, 1980; Stone, 1983).

In fact, the 1980 Army Family Symposium held in Washington, produced major recommendations to improve the quality of military family life, including the establishment of job centers, a world-wide skill bank system, and career planning programs for family members (The Sunday Register, 1982; O'Beirne, 1982).

In response to this, the Department of the Army issued a directive requesting installation Civilian Personnel Officers consider establishment of an employment assistance program to meet the needs of military family members who desire or need to work and whose career or educational goals are disrupted by Army moves. This directive, in addition to suggestions made by the local Fort Monmouth Family Action Committee, and with the commanding General's approval, led to the establishment of a pilot program to assist family members whose sponsors were assigned to Fort Monmouth. In February 1983, the Military Family Member Employment Program became operative. The objective of the program is to provide comprehensive employment counseling and assistance to military family members and to provide a central point of contact for Federal and Non-appropriated fund job information.

Therefore, the intent of this researcher is to present findings that will not only add a new dimension to the current state of knowledge about military spouses, but will provide the Military Family Member Employment Coordinators at Fort Monmouth with an expanded profile analysis of the civilian spouses they counsel especially regarding the importance these military wives place on work-related values and their commitment to a paid job/career.

Methodology

Sample

The subjects in this study consisted of civilian spouses of military members assigned to Fort Monmouth Army Base in New Jersey. Data was collected by a questionnaire, accompanied by a transmittal letter signed by the researcher, distributed through the auspices of the Military Family Member Employment Program office to military wives at career planning workshops, as well as other locations on Base including the Army Community Services Office. In each case, participants were recruited from the individuals in attendance and were reminded of the voluntary nature of their participation, as well as the anonymity of their responses. A total of 265 questionnaires were distributed. The completed questionnaires were returned by U.S. Mail in self-addressed envelopes provided with the questionnaires. A total of 75 usable questionnaires were returned for analysis, representing a response rate of 28 percent. The sample was limited to those individuals who elected to complete the questionnaire. Thus, while there may be some self-selection bias in this sample, such a problem is associated with all surveys where individuals choose whether or not to respond to the questionnaire (Emory, 1980).

Research Instrument

The questionnaire utilized for this study was primarily designed to elicit demographic characteristics and respondent attitudes toward work-related values. The demographic information included age, marital status, parental status, educational attainment, current employment information, as well as spouse's branch and rank of service.

The six work values under investigation were identifiable psychological needs. These values have been defined by Eyde (1962) as the expressed need for: 'dominance/recognition,' 'economic,' 'independence,' 'interesting activity,' 'mastery achievement' and 'social.' The current study utilized the original Work Values Scale reported in earlier research by Eyde (1962) and later expanded upon in a study by Wolfe (1969). The scale contained 84 statements broken down into 14 sets; each set consisting of a work value item from each of the six specific factors. The respondents rank-ordered their responses in each set. Ratings range from 1 (the reason which best applies) to 6 (the reason which least applies).

Analysis of Data

The ranking scores made by all respondents were combined to calculate total scores for the six work values. Mean scores and standard deviations for the values were computed for each variable. The six work values were used as dependent variables and the demographic variables were used as the independent variables. A test for a measure of association was employed in a crosstabulation of selected independent variables and the dependent variables.

Because of the exploratory nature of this research, no *a priori* hypotheses are formally articulated.

Results and Discussion

Personal Data

Demographic characteristics derived from the analysis of questionnaire data show that the respondents range in age from 20 to 50, with a mean range of 30-34. Fifty percent of the respondents report a college or advanced degree in such majors as the Social Sciences, Nursing, Business and Education. Of the sample surveyed (N=75), 95 percent report presently married, with a majority (58 percent), indicating length of time married from 1-10 years. The military rank of service of the spouses reportedly range from Private Enlisted (E-2) to Colonel, with rank of Captain, U.S. Army, reporting the larger number (17 percent). Eighty-eight percent of the military members are with the U.S. Army, with 12 percent reporting other military services including the U.S.N., U.S.A.F. and U.S.M.C. Of the 76 percent reported having children, 57 percent have one to two children, while 19 percent have three or more. Interestingly enough, a recent study by the Bureau of Labor Statistics reported that 75 percent of military families included children, and three out of five of these families had one or more children under age 6 (Jones, 1982).

Work Experience

According to Allyson S. Grossman, an economist with the Bureau of Labor Statistics, 50 percent of military wives are in the labor force (Jones, 1982). This study revealed that 33 percent are working full time, having accumulated an average of seven years in the workforce, while 28 percent are employed part time with an average of four years. Of the number of hours reported working per week for those employed, 41 percent recorded between 34-40 hours/week, and 22 percent recorded 10-30 hours/week. Employed respondents reported an average annual salary range of $10,000 to $15,000. Interestingly, the Census data show that full-time working wives tend to have earnings concentrated in a small range, from $9,000 to $17,000 a year, regardless of their husband's annual salary (Pear, 1983).

Occupational classifications range from professional/technical (43 percent), secretarial (23 percent), sales/clerical (8 percent) and managerial/administrative (7 percent); the balance (19 percent) reporting crafts and other services. As for supervisory experience, 39 percent of the respondents reported from one to five years, while 57 percent reported less than one year. Only four percent reported six or more years. As to the type of company/firm, this study revealed 35 percent are employed by the Government, while 28 percent are employed by service-oriented firms, and 34 percent by nonprofit organizations. Non reporting accounted for the remaining three percent.

Work Values

Figure 1 is a bar graph of confidence intervals constructed for mean responses for statements pertaining to the six work values. The study reveals the respondents demonstrated a high need for their work to yield the mastery-achievement value --accomplishing things; doing a job well and completely. A survey conducted by the Public Agenda Foundation, a New York non-profit organization headed by Daniel Yankelovich which reveals that American workers have a surprisingly high level of job commitment and job effort, corroborates the findings in this study that women in particular are more likely to believe in the work ethic and that they have the "inner need to do the very best I can, regardless of pay..". Women's upbringing may be responsible for their high work ethic. Also, judging by how well women tend to be paid, they would seem to be a first-class buy in the job market (Kagan, 1983).

An overlapping pattern appears to exist with the remaining work values, namely in descending order: social, interesting activity, independence, dominance/recognition, and economic. As for social needs, these women expressed their desire to make new friends, to be around others, and to seek social outlets. The interesting activity value is closely tied into their expressed desire to find social outlets. These military wives seek an opportunity to vary their experience, to avoid boredom and to experience different events. Independence was only a moderate need. The respondents are not basically going to work to free themselves from felt restraints; but there is some need to demonstrate that they could be independent if circumstances should require. This finding appears to be in line with the fact that military wives learn how to be flexible because of their lifestyle with frequent moves and periods of separation that leaves them solely in charge of home and family (Jones, 1983).

Money and power apparently are not what makes these military spouses feel good about their jobs. The Counselors should be aware of the fact that it will be an unusual woman who feels the need to dominate others on the job. It is curious that the military wives surveyed perceive as least important economic rewards of work. At least with this population, Counselors should note this factor and emphasize other values rather than economic. Although, in reality two-thirds of the 47 million women in the labor force must work to support themselves and their families (Mahoney, 1983; Linden, 1981).

Figures 2 through 5 illustrates in line graph format the results of testing for a measure of association between selected demographic variables (age, parental status, educational attainment, occupational classification) and the six work values. Examining the mean scores, revealed the following results. All women in the study regardless of age, parental status, education, or occupation, placed the greatest emphasis upon the mastery achievement value. Therefore, this must be considered as the central work value for these military

spouses. This means that the Counselors must always keep in mind that these women need to derive a sense of accomplishment and satisfaction from their work.

Following mastery achievement, other interesting highlights from the Profile Analyses revealed that the importance of work providing independence are ranked highest for those military wives without children, and for women who have advanced degrees. The expectation that work would yield social rewards was highest for women in the age range 20-24, holding a high school diploma, and with two children. Women with advanced degrees ranked dominance/recognition highest, as did the professional/technical occupational group. Although the economic value of work was low for all women, the value generally tended to decrease in importance as a woman's education increased. Also, the oldest age group (40 and over) placed the least priority on this value, as did the spouses with no children.

A recent Bureau of Labor Statistics study predicts that more and more military wives will be entering the labor market during the 1980s (Jones, 1982). Recognizing the emergence of military spouses in the workforce, and the fact that employment of spouses significantly affects the retention of military members, this perspective implies that with a more informed insight into this growing segment of the women's labor force by the Coordinators can ultimately assist in accomplishing the overall goal of the Military Family Member Employment Program of improving the quality of military family life.

Due to space limitations, Figures 1 through 5 and References will be furnished upon request.

Isolation Matrix: A Tool for Discovery

Kathleen P. O'Beirne
Industrial College of the Armed Forces

"Natty Bumpo didn't have a wife!" The officers in Elective 246 at the Industrial College of the Armed Forces, "Military Family Studies," reacted to this insight with laughter and then some pensive expressions. They remembered the hero of Cooper's Leatherstocking Tales and other such heroes of American literature who had peopled their boyhood reading. They were isolated men, loners, explorers. The women were not complex characters. They were cardboard creatures named for virtues like Honor and Chastity, even though prototypes for heroines abounded in remarkable frontier women.

Today's Natty Bumpos, in the persons of American military personnel, do have spouses who share their potential for heroic stature. Opportunities for meeting challenges above and beyond the commonplace do exist for military family members. Their awareness of the noble dimensions that are possible in their lives can be an important coping skill. Knowledge of their strengths and weaknesses enables military families and their care providers to deal with their omnipresent challenge: isolation.

Recognition of the positive, as well as the negative aspects of isolation can be aided by the following matrix. The five families of which military family members may be a part are:

- Nuclear Family -- one's spouse and children, if married; or one's parents and siblings, if single.
- Extended Family -- one's relatives (who may or may not live close by) and, for a military family, the network of very close friends from one's hometown and from previous duty stations.
- Military Unit Family -- the personnel in the specific unit to which one is attached, and their families.
- Neighborhood Family -- immediate neighbors plus friends in the community in which one lives, on base and off.
- Service Family -- the overall Army or other Service which provides official and unofficial support services to military families.

When an individual has support from four or five of these families, he or she can function well. With only three, he experiences stress. Fewer supports place him at risk.

The vertical axis includes four major types of isolation (that are due to the military setting vs. personality disorders):

- Geographic -- physical separation
- Social -- at the level of acquaintances, surface recognition
- Emotional -- at the level of deep friendships, lasting ties, recognition of one's abilities and thoughts
- Cultural -- a catch-all category which includes:
  - ethnic - foreign born spouses; living overseas or in a new section of our own country
  - aesthetic - visual and performing arts, education, and intellectual endeavors
  - spiritual-philosophical - pursuits of a religious and ethical nature

Isolation Matrix
(c) 9/26/82

Kinds of Isolation	Nuclear Family	Extended Family	Military Unit Family	Neighborhood/ Community Family	Service Family
Geographic					
Social					
Emotional					
Cultural					

*For practical use, this matrix must be much enlarged.

The Isolation Matrix can be used by a wide variety of participants. The officers and spouses who are students in my ICAF course have found that discussion of the matrix allows them to share insights that would be too personal in other contexts, and it accentuates the complexity of the military family lifestyle. An individual may fill in his or her own profile, spotting strengths and weaknesses. A military unit group, such as the Key Wives in the Cherry Point Marine Corps Air Station Family Readiness Program, can work up a generalized profile which will enable them to pinpoint ways in which they can provide support. They will also find areas which the Service Family or one of the other five families could serve best. Service providers, such as family service/support centers, could do likewise.

Some of the examples of isolation that have been suggested by ICAF students and Key Wives may trigger your own thoughts. The symbols (+), (-), and (+/-) indicate the evaluation they placed on each experience. Some "cures" are added.

Geographic:

- o Husband deployed or TDY; physically absent (-), but if communication is possible (+/-).
- o Lack of phone, newspapers for junior enlisted (-).
- o Living at distance from extended family:
  (-) if family is supportive;
  (+) if one joined military to escape bad situation at home
  (+) if one learns independence.
- o Living at distance from other unit families (-).
- o Lack of sponsor (-)
- o Deployed units leave families behind (-). Cures: visit or tour similar ship; use videotapes, films to show activities at home and aboard ship.
- o Ghetto of military housing overseas (-/+).
- o Lack of transportation from base or trailer park to town (-).
- o Fishbowl existence when living on post, no privacy (-).
- o Loss of vistas of one's youth; lack of natural ties with the earth (-). Cures: annual visits "home;" canoeing, hiking, gardening wherever one lives.

Social:

- "Caging" of wife, association with neighbors or unit wives forbidden (-).
- Very young children, especially if no $ for child care (-).
- Lack of skills in making friends, especially true of those who grew up in small town (-).
- Cabin fever, often climatically induced (-).
- Difference of lifestyles (between member and extended family, unit families, or neighborhood families) (-/+). Comment: military families often outgrow a provincial background, come to value differences.
- Strident separation of officer and enlisted families (-).
- Failure by Commanding Officer (CO) to initiate whole unit activities which would help families identify with the workplace, mission, and each other (-).
- Isolation of very senior and very junior personnel unless seniors initiate and assist (-).
- Unit family is by chance, not by choice (-).
- If unit is large, anonymity may result (-).
- Military families seen as lower class by locals (-).
- Social isolation of wife when husband is deployed (-).
- Increase in number of working wives leaves less time for socializing outside of family unit (-/+).
- Sense of immediate common bond with other military families (+).

Emotional:

- Companionate marriage style strengthens a couple, but may isolate them from natural support groups (+/-).
- Reliance on nuclear family due to frequent moves (+/-).
- Some military personnel of the lone hero variety exclude spouses and children from their world (-).
- Differentiate between creative solitude (+) and loneliness (-).
- Lack of continuity in location, hours of duty, roles, etc. requires great emotional flexibility. An individual may grow to this challenge (+), or may decide not to invest in order to avoid hurt upon separation (-).

- Lack of family or close friends nearby to validate individual (-).
- Loss of unconditional love of grandparents and cousins due to distance (-).
- Development of intense friendships with unit families due to common experience (+).
- Lack of knowledge of normal emotional rhythms common to given military assignments (e.g. submarine wives' syndrome) (-).
- Failure by unit CO to value family's role in readiness and retention (-).
- Refusal by some to be more than sociable (i.e. will not risk vulnerability, admission of hardships because it might hurt career) (-).
- Failure to know immediate neighbors (-).
- Strong ties to past unit or location may preclude investment at new post (-).

- o Lack of frequent contact with other military families (-).
- o Lower rank families feel like outsiders in the military community (-).
- o Co-location of military family support services (+)

Cultural:

- o Reliance on television families as "knowns" wherever one moves vs. active cultivation of new friends (-).
- o Continuation of one's own traditions and artistic pursuits as link with past and future (+).
- o Double culture shock for foreign-born spouses to U. S. military community (-).
- o Culture shock of living overseas and re-entry (-/+).
- o Loss of common class ethics or behavior, sense of being different from others (teens, whether spouses or children, feel this intensely) (-/+).
- o Difference in cultural background (racial, ethnic, or educational) of neighbors or military unit families (-/+).
- o Change in living conditions from that of one's past (rural ⟷ urban; density of housing, climate, location) (-/+).
- o Lack of representation for military families on governing bodies (e.g. School Boards) (-).
- o Lack of cultural opportunities for participation or attendance (-).
- o Lack of understanding of military family traditions (-).
- o Lack of religious faith (-).
- o Non-valuing of military profession by others (-).

The predominance of negative factors is common to a first listing; problem areas are more rapidly identified than positives or "cures." The next step in the use of the matrix is to focus on a specific area, looking for positives that were overlooked in the first general delineation.

For example, at pre-deployment briefings one might hear a wife say, "Deployments give me time to really concentrate on my job or to pursue some long-term projects." A husband might admit that he enjoys his time at sea, and though he misses his family, he does not experience the daily friction between his two commitments. A wife might recognize that her self-sufficiency grows when she is responsible for finances, repairs, and child-rearing. Both might see that the deployment is a growth opportunity for all members of the family, an experience that their civilian counterparts probably do not have.

Where negatives remain, the task is to think of preventives or cures. One of the most productive exercises for pre-deployment briefings is to brainstorm ways to keep the person leaving emotionally present even if physically absent. This problem-solving can take place on all levels, from the personal family level to the military unit and Service levels. Although many suggestions that have grown out of such meetings could be shared here, part of the value is the ownership generated by coming up with solutions tailored to the needs of one's own family, unit, or command.

Military families can live lives with heroic dimensions. They explore new domains, both geographic and emotional. They are often alone. Their terrain is full of peaks and pits with very few plateaus. The military lifestyle offers the opposing possibilities of greater success or greater failure than are common to most of their civilian peers. The environment most conducive to human development is one sufficiently changeable to pose constant challenges, but not so severe as to prevent successful response. Use of the Isolation Matrix will help families and their care providers meet the requirements for Arnold Toynbee's ideal environment.

AD-P003 327

A PILOT TEST OF THE USE OF TELEPHONE METHODS TO SURVEY ACTIVE ARMY PERSONNEL

Authors: Ms. Donna Johnson, Major Scott Berry, Department of the Army Soldier Support Center, Dr. Mary Beth Merrin, WESTAT.

Abstract

This paper reports on the results of the first pilot international telephone survey of Army personnel. The objective of the pilot survey was to assess the feasibility of using telephone methods for surveying active Army personnel. The results represent a new development in the field of survey methodology.

In the civilian sector, telephone methods have offered numerous advantages over in-person and mail surveys, including shorter time periods for data collection and improved quality control over the interview process. Despite these benefits, the feasibility of conducting telephone surveys in the military community has been unclear.

This paper discusses the methodology that was developed for conducting a telephone survey of Army personnel and explains how it was implemented. The methodology is then evaluated using several criteria. The results indicate that telephone methods offer a viable alternative to more traditional methods commonly used to survey military personnel.

Background

Telephone methods have become increasingly popular in the civilian sector as a means to conduct surveys of various populations. Telephone methods have offered numerous advantages over in-person and mail surveys, including shorter time periods for data collection and improved quality control over the inter-view process. Despite the growing popularity of telephone methods and the advantages they offer, their utility for surveying active Army personnel had not been assessed prior to the conduct of the work discussed in this paper.

Several situations existed within the unique Army environment that made the use of telephone surveys highly questionable. First, Army troops are scattered world wide. To locate and interview soldiers who are so dispersed and often mobile is no small challenge. Second, the Army has no centralized list of soldiers' home telephone numbers. Conducting a telephone survey without such a list of telephone numbers was even more of a challenge. Third, most of the unmarried soldiers and soldiers in lower grades do not have home telephones because they live in barracks. For these soldiers, conducting a survey using barracks telephones would have been unacceptable.

The unique problems posed by the Army environment made it impossible to simply transfer the established methods used to conduct surveys in the civilian sector to a survey of Army service members. Therefore, a new methodology had to be developed. The methodology was developed in a joint venture by the Attitude and Opinion Survey Branch of the U.S. Army Soldier Support Center and WESTAT, a civilian survey research firm. Details of this methodology and how it responded to the distinct challenges posed by the Army environment are discussed in the following paragraphs.

The Methodology

As we undertook the study, our problems compounded. The topic could not have been more difficult. In addition to the challenges facing any telephone survey of Army personnel, the subject matter of this first pilot project posed additional complications. The subject was the Non-combatent Evacuation Operation (NEO) policy and Family Care Plans. The NEO policy concerns the removal of family members from overseas areas in the event of a mobilization. Family Care Plans concern the movement of dependents of sole parents to the NEO site overseas, or in the continental United States (CONUS) to the home of a designated caretaker.

This subject matter had dual implications for the survey. For one thing, it required that the survey extend to units overseas as well as to those in CONUS. In addition, the subject matter suggested a unit sampling approach so that soldiers and their unit chain of command be surveyed.

The overseas requirement prompted a decision to survey units in the major command (MACOM) of the United States Army Europe (USAREUR). In order to survey units in USAREUR, we considered various options. These options include using:

1. AUTOVON lines;
2. Commercial lines with an 800 toll-free number;
3. Commercial lines with a collect telephone call; and
4. Commercial lines with an appointment call.

AUTOVON Lines

One of the most obvious approaches to interviewing soldiers in USAREUR involved the use of AUTOVON which serves most major CONUS installations and limited overseas areas. However, there were a number of problems associated with using AUTOVON for telephone interviews. In addition to getting DOD approval to allow a contractor to use AUTOVON lines, the largest drawback centered on the poor quality of phone connections between CONUS and USAREUR. Over-crowded circuits, frequent preemption of telephone calls, and echoes in the background were some of the most blatent faults.

Because the use of AUTOVON lines between CONUS and USAREUR seemed impractical, the use of AUTOVON lines within the USAREUR network was considered. However, there were additional problems with this approach. Most notably, no space could be found on the installations in USAREUR for setting up an interviewing facility with adequate privacy and a sufficient number of telephone lines. Therefore, the AUTOVON approach was quickly abandoned.

Commercial Lines with an 800 Number

One of the methodological concepts widely used in the civilian sector which we had hoped to test in the study was the use of a toll-free 800 number. While this approach posed no significant hurdle for CONUS soldiers, we learned that this methodology was not feasible for our European based soldiers. It is impossible for soldiers stationed in USAREUR to either direct dial an 800 toll-free number or use a European operator to place the call.

Commercial Lines with a Collect Call

Similarly, while it is possible for soldiers to place collect calls to the United States from European commercial phones, this method was rejected because these lines are not located in areas convenient for the majority of soldiers stationed in Europe. To ask soldiers to walk/drive to a location either on or off post to place a long distance telephone call would have invited a dismal response rate.

Commercial Lines with an Appointment Call

While the use of commercial lines for soldiers to call in to an interviewing facility had been rejected, the option of using commercial lines for outgoing calls to soldiers from an

interviewing facility appeared more promising.[1] However, it was not sufficient just to attempt telephone calls to soldiers. A mechanism had to be developed which would assure that soldiers would be in a location to receive these in-coming calls and be interviewed. This problem was solved by the use of an Appointment Call Method. This method employed a network of installation points of contact (POC's). The Appointment Call Method together with the POC network became central features of the survey methodology, both in CONUS as well as in USAREUR.

To facilitate coordination with the units to be sampled, major commands were tasked to provide one POC per installation. WESTAT used the POC's to inform the installations of all the details concerning the survey and to obtain sampled unit commanders' and first sergeants' names and telephone numbers. Additionally, the names of sole parents were passed to each POC for confirmation by the unit for the purpose of setting up a scheduled interview time in the unit orderly room. Soldiers were then instructed by their unit chain of command to be in the orderly room at the prearranged time to participate in the interview. Once coordination was completed, WESTAT commenced the telephone interviews.

Since we were testing the feasibility of conducting telephone surveys in the Army, a decision was made to test an alternate method of data collection. This method also involved the use of commercial lines and POC's. In addition, it entailed the use of a toll-free 800 number. This approach was limited to a random subsample of units in CONUS. WESTAT provided POC's with information concerning who was to call, when, and the 800 number. The unit chain of command informed the selected soldiers to call WESTAT during the time frame specified.

[1]The option of using a separate and established contracting firm located overseas was also considered. However, a decision was made not to use a separate firm to conduct the overseas portion of the survey. This decision was prompted, in part, by the objective to conduct all data collection from one central location where rigorous quality control would be executed and calls could be monitored by SSC-NCR and WESTAT staff. In addition, the same problems facing the use of collect telephone calls applied here. That is, soldiers could find it inconvenient to place a call from a German telephone line.

Survey Results and Recommendations

By standards commonly used to assess the adequacy of survey efforts, the methodology, as developed and implemented, was a success. The methodology was to facilitate telephone contact with designated Army personnel and secure their participation in the survey. The major indication that this objective was attained is the study's overall response rate, 62.8 percent. The response rate for the major method of data collection (Appointment Call) was 78.6 percent. The response rate for the 800 Number Call-In Method was 32.5 percent. While the response rate achieved by the 800 Number Method was low, it is slightly higher than our estimate of the response rate for a mail survey with the same topic and population. While there were some problems in getting through to units overseas (due to circuit congestion), the response rate of USAREUR was comparable to the response rate achieved in CONUS (76.8 percent vs 81.1 percent).

The actual data collection period spanned an interval of 16 days. However, the entire survey process, from the initial development stage through data collection and analysis, took approximately 5 months. Had the survey been conducted using traditional mail procedures, the total time frame would have been slightly longer. For future efforts, however, the time period for conducting a telephone survey could be considerably shortened. Currently, the Army is engaged in testing a revised methodology which is to reduce turnaround time considerably.

According to interviewers' assessments, the quality of respondent participation was quite high. Respondents' general attitudes towards the survey were overwhelmingly positive and most demonstrated a very good understanding of the questions asked.

Summary

In summary, telephone surveys appear to be a viable way of surveying Army personnel, both in overseas as well as in CONUS. Moreover, telephone survey techniques offer the Army improvements over other survey methods in several respects. The quality control achieved by supervisory staff (as well as SSC-NCR staff themselves) who silently monitored the interview process is noteworthy.

The Army might have been in the Dark Ages in terms of telephone surveys in 1982. But today we are enlightened.

AD-P003 328

Linkages Between Family Support Variables and Military Career Commitments

Dennis K. Orthner & Joe F. Pittman, Jr.
University of Georgia

Abstract

This study is designed (1) to test an empirical model which describes the linkages between the family and community variables which contribute to Air Force member job commitments and spouse support for members' Air Force careers and (2) to identify the preliminary impacts of the Family Support Centers on the variables in the model. The data were collected from probability samples of Air Force members and spouses and analyzed by means of path analyses. The results indicate that family support variables contribute to over one-third of the job commitments of members and one-fourth of the support provided by Air Force spouses.

Introduction

Study Objectives

This study has been designed to test several of the key assumptions underlying the U.S. Air Force Family Support (FSC) program. These assumptions predict that (1) family and community factors play a role in Air Force member job and career commitments and that (2) improvements in programs and services for families will result in increased family support for the Air Force and increased job morale and commitments of Air Force members.

Linkage Between Family and Job Commitments

Interest in the factors that contribute to job performance, morale and commitment has been growing. A great deal of attention in the research and human resource literature has been given to the value of training, management and job enrichment for human productivity. The importance of organizational support and of psychological well-being for job commitment and performance is well established. This has led to the development of many high quality personnel and human resource programs to improve management competence, work environmental conditions, and personnel support systems.

The potential importance of family factors to work related commitments has been more recently suggested. In part, the reason for incorporating family considerations into occupational support systems has been based on management studies which have indicated that family factors are increasing in importance among those who are making career decisions. They also appear to be influencing the level of stress or well-being which men and women bring to their job. In addition, demographic trends indicate that the composition of the family has changed from the traditional model in which the wife and children were at home supporting the working husband to many models which incorporate employed wives, dual-career marriages, childless couples, and single parents. These newer family forms sometimes require

greater investment on the part of the worker and can add new strains which have to be accomodated in the workplace.

The present study is part of a larger Air Force sponsored evaluation of the FSC program. The overall project from which this study is a part is designed to assess the high priority needs of Air Force families and determine the impact that FSC's have on meeting those needs, on improving the quality of life on Air Force bases, and on increasing the effectiveness and commitment of Air Force personnel.

The overall objective of the study is to prepare and test an empirical model which describes the contribution of family and community variables to job and career outcomes which contribute to the Air Force mission. The components of the model have been selected because of their direct and indirect linkage to FSC programs and services. The model, therefore, provides a statistical prediction of the potential influences that the FSC program can have on the job commitment of Air Force members.

Methodology

Sampling Procedure

The data were collected from probability samples of Air Force members and spouses by means of mail surveys conducted in April, 1982. The data were originally collected as part of a baseline survey designed to devise guidelines and offer recommendations for the planning, development, and operation of Air Force Family Support Centers. The surveys included measures of the knowledge and utilization of and satisfaction with current programs and services; methods of coping with personal and family problems; and the quality of individual and family life.

The surveys were conducted at nine Air Force installations. Five of the installations had prototype Family Support Centers. Four matched installations that did not have Family Support Centers were selected and surveyed as control sites. The target sample at each base consisted of 375 married Air Force members and 375 civilian spouses. The mail survey was implemented following a four-wave method designed to obtain a maximum number of responses. Overall, the response rates were a respectable 80% for married members and 72% for spouses.

Not all of the data collected were utilized in the present investigation. Since the focus of this study is on family influences on Air Force member job and career attitudes as well as the impact of the FSC on the factors that contribute to those attitudes, a decision was made to reduce the sample size in order to test our hypotheses with the people most relevant to the issues under study. First of all, only six of the nine bases are included in the analyses, These included the three most operational FSC bases and their control bases. While these three FSC programs had only been operational for four to six months, they were beginning to generate impacts on the people in their base community. The other two FSC programs were not yet operational at the time of the study. The sample was also restricted to include only those Air Force members with less than ten years of service. This optimized the variance in job morale and commitment, the dependent variable in the study.

As a result of these restrictions, the final sample upon which the data

analyses were conducted includes Air Force members and spouses with less than ten years service on six Air Force bases, three with FSCs and three without FSCs. The final samples include over 1300 persons, 60% of which are members and 40% of which are spouses. This sample size is more than adequate to statistically test the path models and hypotheses developed for the study.

The data were analyzed by means of multiple-regression, path analysis techniques. Two separate path models were developed, one for members and one for spouses. The member model included factors predicting Air Force job commitments while the spouse model included factors predicting support for their spouses Air Force career. Together, these two models converge into a composite picture of the family and community variables that statistically contribute to personnel support for the Air Force mission.

Hypotheses

A primary objective of this study is to evaluate the impact of operational Family Support Centers on the attitudes and behaviors of Air Force members and spouses. In addition to the primary objective, several other goals are considered important. For Air Force members, these goals include: (a) assessing the utility of personal variables, familial variables, and program variables in predicting the commitment of members to their Air Force job; and (b) studying the interrelations of these personal, familial and program variables. For Air Force spouses, the goals are similar. Rather than predicting job commitment, however, the focus is on the prediction of support for the member's job commitment.

In order to meet the above objectives, two theoretical models were developed. These models illustrate the theoretical underpinnings of the more detailed models that are developed and tested with the path analytic procedure.

The theoretical models indicate that personal and familial variables are hypothesized to be important predictors of job commitment and spousal support for this commitment. Furthermore, it is hypothesized that these two clusters of variables will be strongly interrelated. Another cluster of variables, the program variables, is expected to be interrelated and a relevant predictor of the personal and familial variables, but not a direct predictor of career commitment. Finally, the FSC variables are expected to have their greatest influence on attitudes and behaviors toward established programs. Therefore, the FSC is expected to affect the quality of Air Force life through its effects on the perceptions of, knowledge about, use of, and satisfaction with established base programs. Furthermore, these FSC variables are expected to be affected, in a reciprocal fashion, by the program variables that they influence.

Data Analysis

In order to test the two models, fourteen variables were incorporated into two separate, detailed models. The first model was developed for the Air Force members, while the second was developed for Air Force spouses. A path analysis procedure was then run on each model. For Air Force members, the dependent criterion variable of job commitment composed of several questions regarding job satisfaction, morale and career commitment. For spouses, the dependent/criterion variable of spouse support is composed of questions regarding the spouse's willingness to support the Air Force career of her husband.

Findings

With a few exceptions, the results of the data analysis confirmed our hypotheses. For Air Force members, the variables included in the path model explained more than one-third of the variance in the dependent variable (R^2 = .3550). The strongest predictors of job and career commitment came from two family variables, i.e. spouse support (B = .288) and the feeling that the Air Force is a good place to rear children (B = .171); and a personal variable psychological adjustment (B = .280). We did not expect any general program variables to directly impact on the job and career commitment of Air Force members but one variable did have a significant impact: the feeling that the base was responsive to family needs (B = .175). This indicates that base responsiveness to families is a more important contributor to job morale and commitment than we had earlier hypothesized. It also suggests that Air Force programs need to impact on the overall quality of life on the base, not just on the immediate personal and family concerns of people in the base community. Apparently, the overall perception that the Air Force cares for its families is very important to married Air Force members.

There are other factors in our statistical model that also impact on career commitments, but they do so in an indirect manner. For example, marital satisfaction does not directly impact on career commitment but it does have a strong impact on the personal adjustment and well-being of Air Force members and also on the feelings of these members that their spouse is supportive of their Air Force career. This means that improvements in the quality of Air Force marriages will have very important indirect effects on the job related attitudes of Air Force personnel. Likewise, improvements in marital satisfaction have a direct impact on the size and quality of the support network of Air Force members. This, in turn, increases their feeling that the Air Force is a good place to rear children which, in turn, increases their spouse support and their own support for the Air Force. Therefore, even though improvements in marital quality are not directly related to job and career commitments, they do have a strong indirect effect through other personal and family factors which subsequently impact on the job commitments of Air Force members. This "domino effect" of the variables substantially improves our understanding of how human service programs can ultimately have an impact on the mission of the service.

Family and personal adjustment factors were also found to be very important to the support spouses give to the careers of married Air Force members. Nearly one-fourth of the variance in spouse support was accounted for in the path model (R^2 = .2214). The feeling that the Air Force is a good place to rear a child plays a major role in spouse support (B = .342). The personal well-being of the civilian spouse is also a significant predictor of her support (B = .176). Two program related variables contributed significantly to spouse support as well: the feeling that the base was responsive to families (B = .081) and satisfactory previous experiences with Air Force programs (.091). This latter finding suggests that improvements in programs for families and improvements in the overall quality of life for families are likely to continue to have a significant direct impact on spouse support.

Given the fact that the Family Support Centers at these installations were only beginning to become operational, it is not surprising to find that the presence of these programs had little direct effect on either spouse support or job and career commitments at the time the data were collected. In fact, it is surprising that any statistically significant differences appeared between the

FSC and control bases after only a few months of operation. Still, as hypothesized, the Centers were making an impact on several of the program related variables in the model. The data indicated that Air Force members and spouses at FSC bases were significantly more knowledgeable about family programs and services and they were more likely to feel that their base is responsive to family needs. These findings are significant since they indicate a predictable pattern in FSC developments and potential for future impacts through other intermediate variables in the model.

Conclusions

The findings indicate that there are both direct and indirect ways through which the Air Force impacts on family attitudes and the job morale and commitments of its personnel. First of all, the Air Force directly impacts on its personnel through its policies toward families. These are interpreted by many Air Force members as supporting their ability to provide for their families. This path of influence from the Air Force to the member is especially important in the data. Since most married personnel are men and their definition of themselves as providers for their family is important to them it is not surprising that their belief that the Air Force backs them up in this role results in greater commitment to their job and career. Indirectly, the Air Force has another rather strong capability to influence members' commitments, this time through the spouses and children. The findings from this study are very clear about the importance that married members give to the attitudes of their family toward the quality of life in the Air Force. Members who believe that their children and their spouses are adjusting well to the Air Force are much more likely to indicate a desire to stay in the service and more willing to report that their job morale and performance is high. The data from spouses supports this chain of influence. The spouses themselves indicate that their support for their husband is very much a product of their belief that the Air Force is supporting them and their family. When that support is defined as weak, then the spouse is less willing to support her husband's career.

PANEL SESSION

PANEL SESSION: BIOFEEDBACK FOR ASSESSING AND DEVELOPING SELF CONTROL

SESSION CHAIR: Russell N. Cassel (Cassel Psych Center)

PARTICIPANTS: Russell N. Cassel (Cassel Psych Center)

Joseph Kamiya (Langley Porter Institute)

Mark S. Schwartz (May Clinic)

Patricia Cowings (NASA-Ames Research Center)

PROCEEDINGS ENTRIES

"Biofeedback for assessing and developing self control" (Summary of program and introduction)

"Biofeedback: What, why, when, for whom and by whom?" (Mark S. Schwartz)

"Autogenic feedback training for astronauts: The space adaptation syndrome" (Patricia S. Cowings)

"The electrogastrogram as an index of motion sickness" (Joe Kamiya)

BIOFEEDBACK FOR ASSESSING AND DEVELOPING SELF CONTROL (BIOFED)

Panel Members

Russell N. Cassel, Colonel, Air Force Retired.(Chairperson)
Editor & Publisher, PROJECT INNOVATION,
Director, CASSEL PSYCH CENTER,
A Computerized Biofeedback Clinic,
717 3rd Avenue,
Chula Vista, California 92010

Joseph Kamiya, M.D. (Psychiatrist)
LANGLEY PORTER INSTITUTE,
401 Parnassus,
San Francisco, Caliifornia 94143

Mark S. Schwartz, Ph.D.
MAYO CLINIC,
ROCHESTER, MINNESOTA 55901

Patricia Cowings, Ph.D.
NASA AMES RESEARCH CENTER,
Mail Stop N239A-2,
Moffett Field, California 94035

Introduction

Recent research suggests that biofeedback may have important implications for military and other purposes. Jack Sorenson, Educational Director, Civil Air Patrol, Montgomery AFB, Alabama, recently released a study using the electroencephalograph (EEG) to assess the learning effectiveness of more than 6,000 elementary youngsters. These findings may be pertinent to technical training planning. Dr. David Tiedman, formerly with Harvard University, is using EEG as a measure of wakefulness, and for career planning purposes. This may have important implications for pilot selection, and idenfication of accident prone individuals. The Cassel Psych Center has been experimenting with need gratification assessment for both the dominant and non dominant sides of brain. This may have important implications for dealing with personal adjustment problems of people. Typically, the use of biofeedback utilizes a biofeedback loop for the development of self control of select aspects of behavior. Dr. Elmer Green, Menninger Foundation, Topeka, Kansas suggests that such feedback loop is experienced based, utilizing nondominant brain. Because the dominant brain exercises veto power over the nondominant brain, it is crucial that techniques be used to involve the dominant side of brain in use of biofeedback. This has important implications for all experienced based learning. The three discussant members will deal with aspects of biofeedback pertinent to their own particular interests.

BIOFEEDBACK: WHAT, WHY, WHEN, FOR WHOM AND BY WHOM

Mark S. Schwartz PhD
Mayo Clinic and Mayo Foundation
Rochester, Minnesota 55905

Both experimental and clinical evidence strongly support the value of biofeedback (BFB) in the treatment of various health disorders. The BFB field is maturing well and has entered its adolescence. Clinical BFB has become an important part of health care and is used extensively in many major medical settings. Real and "pseudo" issues, disagreements, and skepticism still exist but are diminishing. The purpose here is to provide some of the rationale and considerations for using BFB and to introduce the application of it to the prevention and treatment of Space Adaptation Syndrome. Strictly speaking, BFB is the use of electronic and electromechanical instruments and procedures for accurate measurement, processing, and feeding back, to patients or subjects, physiological information of which they usually or always are unaware. The goal is to enable them to develop some self regulation of their physiological processes.

We all have BFB, i.e., "living feedback" -- already built into our bodies. Much of our internal BFB occurs without our awareness, and ordinarily this is preferable. Better terms for what we refer to as BFB are "augmented proprioception" and "external physiological or psychophysiological feedback." The term "BFB" has been coined to denote these augmented or external feedback systems. Approaches to explaining the process and the outcome of BFB include use of cybernetic, learning, and information-processing models.

In a more general and practical sense, BFB should, and often does, encompass much more than instruments feeding back physiological information. Skillful cognitive preparation, observation, and interviewing and the development of effective therapy and learning paradigms are also very important. Additionally, the skillful application of other physiological self regulatory procedures is most often, but not always, an integral part of therapy. BFB is not a simple procedure. Creativity, flexibility, "clinical" insight, and interpersonal and educational skills and procedures all are increasingly recognized as important.

There are many advantages to using BFB. It facilitates the focus of the patient or subject on the development of self regulation, emphasizes the relationships among thoughts, feelings, and physiological effects, and makes one aware of what is otherwise unavailable to awareness. It is an interesting approach for learning what often is otherwise not learnable or less easily learnable. It permits faster and more reliable learning. It also instills more confidence in the patient or subject and in the therapist or teacher.

Current experimental evidence supports the value of clinical BFB in the treatment of tension, vascular and combination headaches, Raynaud's disease, several neuromuscular disorders, and fecal incontinence. Clinical BFB is also justifiable for selected persons with other disorders such as hypertension, bruxism, anxiety, and functional vomiting as long as appropriate standards and guidelines are used and cost effective procedures are followed (e.g., "stepped care"). When we detect an abnormality in a physiological function that can be measured accurately and reliably, amplified, transformed and meaningfully fed back and whenever these abnormalities are believed to be contributing factors to a disorder and potentionally subject to some degree of voluntary control, I think it is reasonable to consider using BFB.

The cognitive preparation of the patient or subject and compliance also are major factors increasing the likelihood of effective outcomes. How the rationale, the procedures, and the goals are presented and the content of those presentations are important to effectiveness and they influence compliance. Compliance is a complex concept, involving the patient's and the therapist's attitudes and characteristics and their interpersonal behavior. Patients or subjects must understand clearly and must be comfortable with and have confidence in the therapist and the procedures. They also need to have or to develop reasonable responsibility for the achievement of the desired self regulation. These ingredients obviously take us far beyond listening to or looking at one's muscle activity or skin temperature.

The BFB therapist or "teacher" should be familiar with both the physiological and the psychological aspects of BFB procedures in the general usage of the term. Requisite training and experience are obvious. The Biofeedback Certification Institute of America (BCIA) is the national agency for providing credentials in the use of BFB and is fully recognized and approved by the National Commission for Health Certifying Agencies (NCHCA).

The future for BFB research and applications is bright. Despite the costs of instrumentation, the delivery cost need not be significantly different from that of other physiological self regulation therapies used alone. With respect to other cost criteria, BFB can be less expensive when it results in successful outcomes not otherwise as likely, thereby reducing the costs associated with continued symptoms.

It is relevant here to note that physiological self regulatory therapies, including BFB, often have been reported to be effective for functional emesis. The literature and my own clinical experience with many patients have convinced me that some of these procedures can be very useful in eliminating or greatly reducing the frequency and duration of functional emesis. Drs. Patricia Cowing and Joseph Kamiya will speak more authoritatively on the applications to ground motion sickness and the implications for and application to the Space Adaptation Syndrome. There are many facts and guidelines to consider in extrapolating from the experience with ground based applications to the needs of aerospace applications. Briefly, physiological arousal is not highly related to subjective arousal; emotional factors, and the associated physiological arousal, are commonly thought to be involved in motion sickness, including the nausea and vomiting; BFB has been shown to be useful and effective in facilitating physiological self regulation of the autonomic nervous system.

We humans have far more potential to regulate our physiology than we once thought. The frontiers of this field are as vast as are those of outer space. It seems fitting that combining the two fields of exploration and application should be encouraged.

TITLE: AUTOGENIC FEEDBACK TRAINING FOR ASTRONAUTS:
THE SPACE ADAPTATION SYNDROME

Patricia S. Cowings, Ph.D. NASA-Ames Research Center, Moffett Field, Ca.

During space flight about one half of all astronauts suffer symptoms resembling motion sickness. Autogenic-feedback Training has been effective in preventing motion sickness in ground based tests. Selected Space Shuttle crewmembers will be given training with the method to evaluate its effectiveness in space.

Autonomic Changes During Motion Sickness. In seven experiments at the Psychophysiological Reseach Laboratory at NASA-Ames Research Center, physiological recordings of nearly 200 men and women have been taken during motion sickness stimulation (rotating chair and vertical accelerator). The measures taken have included heart rate, respiration, peripheral vasomotor activity, and basal skin resistance. Increased sympathetic activation is generally evident in these measures, but with clear individual differences with respect to both the specific autonomic measures that are most affected by motion sickness, as well as their time course over the duration of motion stress. Motion sickness susceptibility is associated with increases in heart rate in the early phases of motion sickness.

Autogenic-feedback Training as a Treatment for Motion Sickness. Might the symptoms of motion sickness be prevented or diminished by training persons to increase their voluntary control of the autonomic activity associated with such sickness? To answer this question a combination of Autogenic therapy and biofeedback was designed to optimize the training. Half the subjects who participated in the above studies were randomly selected to serve as training subjects, while the other half served as controls. The feedback signals used were analog and digital displays of each of the above physiological functions. Starting with one display for one of the functions, the number of displays was progressively increased until all feedback signals were presented simultaneously. A total of 6 hours of training was administered in tweleve, half-hour sessions conducted on consecutive days. Motion sickness inducing tests were administered before, during and after training sessions to evaluate physiological reactions to this stress. In all of the experiments conducted, there was a clear effect of the training. The experimental subjects showed significantly greater resistance to motion sickness than their controls.

Space Shuttle Experiment. Scheduled to begin this year are two Space Shuttle flights with selected astronauts who will be given Autogenic Feedback Training prior to launch. To be able to evaluate the success of the training as well as provide physiological information concomitant with any of the symptoms of space flight, the crewmembers will wear miniaturized versions of the continuous physiological recording equipment used in the laboratory. Feedback displays will be available, for both daily use to prevent symptoms occurrence as well as to counteract symptoms if they should occur.

TITLE: THE ELECTROGASTROGRAM AS AN INDEX OF MOTION SICKNESS

Joe Kamiya, Ph.D., University of California at San Francisco

Previous work by Cowings and Toscano at NASA-Ames has demonstrated that several measures of autonomic function can be used as indicators of motion sickness, in ground-based tests. To extend these findings the present study, also conducted at NASA-Ames, explored an indirect measure of gastric activity. The electrogastrogram (EGG) of 20 young men and women was continuously monitored during induction of motion sickness with a rotating chair. The electrical potentials were recorded from six bipolar surface DC electrodes placed in a circular pattern over the stomach area, filtered with a band pass of 0.01 Hz. to 0.10 Hz, and subjected to power spectral analysis.

Following resting baselines, the subjects were rotated at an initial speed of 6 rpm for five minutes while they tilted their heads every two seconds to the front, back, left, and right, in random order. The rotation speed was increased by 2 rpm every five minutes until the subject reported extreme maliase. Recording was continued for ten minutes after the rotation stopped.

There was a steady increase in the amplitude of EGG activity starting from the beginning of rotation through the initial recovery period. This was observed in each of four frequency bands of the total spectrum. The higher frequencies showed less amplitude. The specific contribution of stomach activity itself to the potentials can not be evaluated without more direct measures. However, the electrical potentials observed are a promising objective indicator of the development of motion sickness symptomatology. If confirmed in further studies, the EGG could supplement the presently established autonomic indicators.

PANEL SESSION

PANEL SESSION: PERCEPTUAL FACTORS IN LOW ALTITUDE FLIGHT

SESSION CHAIR: Ralph Norman Haber (University of Illinois at Chicago)

PARTICIPANTS: Edward H. Houle (Nellis AFB)

Grant McNaughton (AF Inspection and Safety Center)

Milt Miller (A-7 Fighter Weapons School)

Ralph Norman Haber (University of Illinois at Chicago)

PROCEEDINGS ENTRIES

"Perceptual factors in low altitude flight" (Summary of program and separate paper)

Perceptual Factors in Low Altitude Flight

Program Organizer: Professor Ralph Norman Haber
Department of Psychology
University of Illinois at Chicago
Chicago, Illinois 60680

General Abstract

Mission requirements that compel low altitude fighter operations have become more and more common, dictated by weather, electronic and visual detection avoidance, and the presence of low flying aggressors. Flying low and fast gives no margin for error, and successful performance requires the pilot's full understanding of the interactions of mission task demands, flight aerodynamics, usefulness of available visual information, and allocation of attention. This program summarizes what is known about these factors; how that knowledge is applied to pilot training; and examines areas where further basic and training research is needed. The following topics are covered: what does flying low and fast require of perception and attention; what aspects of the terrain provide correct information for ground clearance and terrain avoidance, as well as target location; how is that information processed by a pilot; how can available visual information be supplemented by useful instrumentation and equipment design; and how do task demands interact with visual perception, especially how the limitations imposed by the aerodynamics of low altitude maneuvering affect attentional and perceptual processes.

Program Participants

Mission Requirements in the Low Altitude Arena.
Edward H. Houle, Capt, USAF; Instructor, F-16 Fighter Weapons School Nellis Air Force Base, Nevada.

Why does a fighter pilot have to fly low and fast? What are the tasks that he has to perform? What are the maneuvers and operations typical of low altitude flying, common to all fighters, as well as those specific to particular aircraft and mission demands? What are the known aerodynamic limitations of aircraft that become critical when flying very low and fast?

Collisions with the Ground--What We Learn from Mishap Analyses.
Grant McNaughton, Col, USAF, MC. Chief, Life Sciences Division, Directorate of Aerospace Safety, Air Force Inspection and Safety Center, Norton Air Force Base, California.

Collisions with the ground while flying low altitude maneuvers is probably the single largest category of mishaps in military aviation, reflecting the special difficulty and risk of operations in this flying environment. Exhaustive analyses of each mishap have helped identify specific factors that contribute to mishaps while flying at low altitude, and have suggested a variety of improvements in training, mission specifications, and aircraft and instrument design. A number of examples are described and discussed to illustrate these problems and suggested improvements, as well as to highlight instances in which we still have problems without clear solutions.

The Visual Information Available for Terrain Avoidance--Can a Pilot See Where He Is and Where He is Going?

Milt Miller, Capt, Arizona Air National Guard. Instructor, A-7 Fighter Weapons School, Arizona Air National Guard, Tucson, Arizona.

To determine the usefulness of perception, it is necessary to know the sources of information available to the visual sense from the terrain, the aircraft instrumentation, and the pilot's knowledge of the particular environment. Such analyses suggest that different terrains provide vastly different amounts of information, of vastly different accuracy and usefulness. Equally important, such analyses show that the different sources of visual stimulation are not equivalent in providing information to the pilot about his altitude above the ground, his range from specific features on the ground or other aircraft, the direction of his flight path, or the rate with which any of these factors are changing. These analyses are presented to show what information is available, and what use can be made of it. One conclusion appears likely; pilots are able to determine their direction of flight fairly accurately, but can get little sense of their elevation above the ground or their range from specific objects. However, there are still great gaps between what is known about the visual perception of pilots, and how pilots can and do perform during low altitude flight. Specific indications for further vision research are labelled, as well as demands for inclusion of visual processing training in low altitude flight training.

The Processing of Vector, Altitude and Range Information during Low Altitude Flight: Interactions with Mission and Task Demands.

Ralph Norman Haber, Professor of Psychology, University of Illinois at Chicago, Chicago, Illinois.

How does a pilot use relative motion changes, or optical flow across peripheral vision, or known size of objects, or texture density changes, or any other visual source of information to perceive his position in space near the ground? Can pilots be trained to attend to relevant sources of information, and to avoid sources that might be mis-informative? Can pilots be trained to know where to look, and what to look for? Can pilots be trained to attend to visual information about ground clearance, even when mission demands require their attention elsewhere? To begin to provide answers to these questions, the visual processing of vector, altitude and range is described, to the limits of our current knowledge. These factors are then integrated into a task analysis that describes the interaction of perceptual and mission demands with the overriding task of avoiding a collision with the ground. Implications for further basic and training research are outlined, drawing on the information contained in each of the presentations in this program.

Perceptual Factors in Low Altitude Flight

Ralph Norman Haber
University of Illinois at Chicago

Abstract. The pilot's number one priority when flying low and fast is avoiding the ground. The ground gives no margin for error. Survival and successful performance requires the pilot's full understanding of the interactions of mission task demands, flight aerodynamics, usefulness of available visual information, and allocation of attention. This paper summarizes what is known about these factors. The following topics are covered: the basic aerodynamic performance characteristics of concern during low altitude flight; what aspects of the terrain and instrumentation provide correct information about ground clearance and terrain avoidance; what flying low and fast requires of perception and attention; and how the interaction of mission task demands with the limitations imposed by the physics of the aerodynamics restrict the conditions of attending and perceiving.

Introduction. A pilot is flying in a low altitude environment whenever his primary task is to avoid hitting the ground. If a pilot can plan his upcoming activities for the next few moments without any concern about the ground, then he is not flying at low altitude. Thus, low altitude is defined by the tasks, not by the numbers. Given the speeds and maneuverability of modern jet fighters, low altitude undoubtedly always includes the 1000 feet closest to the ground, but may extend to 5000 or even 10,000 feet above ground level (AGL) for some maneuvers. This definition stresses on the relative importance of tasks. Low altitude means that the primary task is to avoid colliding with the ground, and all other tasks must be handled during whatever free time the pilot has when he does not have to monitor ground clearance. Free time is what is available after ground clearance is handled, and into free time must go all other tasks, including those pertaining to the mission, and to threats.

Current analyses of accident rates (e.g., Miller, 1983) suggest that low altitude flying is far more dangerous then outside of that enviornment. Miller estimates that the probability of an accident (and fatality) is five hundred percent greater when flying at low altitude than in any other category. Since a high accident rate suggests that other aspects of performance are below standard, it is likely that mission requirements are also sacrificed when flying at low altitude. While flying at low alitude is both difficult and dangerous, it is coming to occupy a a greater and greater percentage of flying time for all types of military aircraft. It is this combination of increased usage, higher accident rate, and presumed poorer performance that dictates the urgency of investments of time and research on problems of low altitude flying. This is the topic of this paper.

I shall briefly describe four components of an analysis of low altitude flying. These include two concerned with physics, and two with psychology: the physics of the aerodynamics of low altitude maneuvers; the physics of information available from the terrain, gravity, and instrumentation; the psychology of extracting and using information to perceive velocity vector, AGL, attitude, and distance range; and the psychology of attention and performance by which mission demands can be met while avoiding the ground.

Aerodynamics--The Physics of Flight. At high altitude, the consequences of a change in AGL resulting from a maneuver are usually unimportant. But at low altitude, the pilot has no margin for error. He must know the physical

consequences of each control input for each maneuver, consequences specified in terms of change in AGL, attitude, speed, and velocity vector.

For example, when flying straight and level at 500 knots, it takes only seven seconds to lose 100 feet altitude if the nose drops just one degree. Recovering from the loss requires some human and aerodynamic reaction time, so available free time (during which the pilot is free to ignore his ground clearance) before inpact is inevitable is substantially less than seven seconds. The pilot has to know these numbers since the aerodynamics of level flight plus the handling characteristics of particular aircraft determine how much time the pilot can spare from the overriding task of terrain avoidance during which he can perform other tasks related to his mission.

As a second example, the most frequent single maneuver made during low altitude flying is a level turn. This requires that the bank angle be exactly compensated by the G forces. If not enough G, or too much bank angle, altitude loss occurs. This descent rate increases with the square of time, so it is not a linear function--twice as much AGL does not give twice as much free time. Since matching bank angle to G cannot be guaranteed by instrumentation or by feel, it is necessary to monitor ground clearance (or an altimeter or its equivalent) to determine if the aircraft is converging with the ground during a turn. Given the exponential function of altitude loss, the amount of free time during turns is very small.

These are only two examples of how altitude, attitude, speed, and velocity vector change with particular maneuvers. Pilot training for low altitude flying must include defining terrain avoidance as the number one task, and must explicitly teach the aerodynamic parameters of each maneuver in terms of the amouunt of available free time it allows away from ground clearance. The physics is unforgiving, and therefore must be paid homage.

<u>The Information About Ground Clearance.</u> Information comes from several sources--through the canopy as reflected light, from instruments inside the cockpit, and from changes in gravitational forces. These are all descriptions in physical terms, since they represent available information, quite separate from how that information is used or interpreted. Such descriptions have not been carried out systematically for the low altitude environment, nor has experimental psychology provided an agreed-upon model of how such descriptions should be specified. Consequently, the demands for an understanding of low altitude flying require development of new ideas, few of which have yet been carefully tested.

It now seems likely that a pilot has no visual information that specifies absolute distance from his aircraft to the terrain, neither that directly below (AGL) nor down range distances to obstacles or targets. A pilot may be able to estimate relative distances, or may be able to perform on-line computations to arrive at an absolute value (e.g., note the amount of time it takes an object to drift between two marks on the canopy when flying at a known speed), but he probably cannot perceive distance in metric units. What he can perceive quite easily is a <u>change</u> in his AGL or range from the information available in the changes produced by relative motion as he flies past objects.

Miller (1983) has provided one analysis of the sources of information about distance and position contained in different kinds of terrain features about distance and relative position. These include the presence of sufficient density of features on the terrain, features that provide the content of optic flow across the retina; the presence of features of known size that could be

used to perceive distance; and the presence of fine details on objects that can be used to determine distance (while cacti come in many sizes, if you can see the thorns, you have to be pretty close!). From this kind of list Miller argues that a pilot should be able to rate each terrain over which he is flying for its informativeness. For example, water, snow, desert and dry lake beds would score quite low on all three of these factors, showing that such terrains provide the pilot with little information from which he can perceive his AGL, range, or any change in AGL, attitude, speed or velocity vector. It is not that some terrains create visual illusions for a pilot: rather some are more informative than others, and it is the underinformative ones that pose special dangers for low altitude crossings.

The different sources described by Miller have very different properties. For example, known size and fine detail must be picked up by foveal vision, requiring visual attention. But optic flow, resulting from the presence of sufficient featural density, probably is informative to the peripheral retina, and does not require an act of visual attention to see. Thus, the different demands on attention by the various sources of information make their value change as the flying tasks change. The permanency of the information also varies. For example, pilots talk of a sense of ground rush when they get very low, but ground rush and probably all aspects of optic flow adapt easily, and are completely dependent on a fixed ground speed for their predictive stability. Another source of visual information comes from the unmasking of a far hill by a near one as the hills are approached. This is a critical cue to whether the near one will be cleared--if there is no uncovering of the far hill, the velocity vector is below the ridge of the near one. This very specific component of motion perspective is perhaps more important than the general case, since it provides a direct answer to an immediate question: will I clear that ridge? We need to know much more about all of the aspects of motion perspective, including motion parallax, to isolate those aspects (such as uncovering) that are informative to a pilot flying at low altitude.

At least at present, nearly everything a pilot needs to know is only available outside of the cockpit. Information from instrumentation in cockpits is not used much. Most of what is available to pilots has a slow respond time, or is not responsive to the most useful aspects of the terrain. Further, pilots are naturally resistent to depend on instrumentation during low altitude maneuvering. The same applies to information from changes in gravitational forces acting on the pilot. Current evidence suggests that most of what a pilot can learn from vestibular inputs is either irrelevant or misleading, as we are now learning from analyses of spatial disorientation.

<u>The Processing of Visual Information</u>. Given the available information about AGL, velocity vector, attitude and speed, how does a pilot use it so he can correctly perceive where he is, and where he is going to be in the next few moments? One common answer is that adult visual perception is automatic and inflexible, so that we do not need to worry about further training. Whatever information is present will be used by a pilot to perceive the scene correctly. This answer is wrong for low altitude flying, for a number of reasons.

(1) A pilot can be taught to assess the terrain over which he is flying for its visual informativeness. Even more important, he can note whether he can depend on optic flow to tell him about unintentional nose slices, or whether he must use foveal vision directed by explicit visual attention to details. Underinformative terrains can be crossed safely if a pilot knows what substitutes must be made. Just as a pilot is taught to switch from visual to instrument rules when appoaching a landing through a black hole, so must he

learn when the different sources of information might be less useful, given the character of the terrain. (2) A pilot might be taught to estimate absolute distance and AGL through practice (with the use of his radar altimeter), so that after practice, he can tell at a glance what 100 feet AGL looks like. Maybe he can learn to use the presence of known sized objects for absolute distance? We do not know the answers to these suppositions, though many pilots claim they have done it themselves. (3) Can we train a pilot to compensate for his peripheral vision losses during high G turns, so that the missing optic flow information can be replaced by a substitute during the turn? (4) Instructors suspect that pilots do not know intuitively how to predict their ground track during turns, nor where to look during turns to determine ground clearance. It is not obvious, and probably does need to be trained. (5) Can we teach a pilot to be suspicious of information that is often misleading? For example, an accelerating descent is usually perceived as a climb, because of the positive G forces, and has accounted for more than one collision with the ground when the pilot thought he was gaining altitude. (6) When can a pilot use a snap glance, as in checking six, and when must he make an extended visual fixation--how long does it take to acquire and to assess visual information of different kinds? We desperately need to know this, because time is in very short supply at low altitude. (7) Current visual selection of pilots stresses their fine detail acuity, and virtually ignores abilities to respond to contrast among low spatial frequencies. Yet we know that many critical flying tasks, such as detecting a distant aircraft, requires the latter ability as much if not more than the former.

This list can obviously be increased. Each item is an example of where training (or selection) might be expected to increase a pilot's ability to perceive his position more accurately or more rapidly. The bottom line in low altitude maneuvering is that avoiding the ground is the number one priority, and to do that the pilot must be continually processing visual information about changes in his position in relation to the ground.

<u>Division of Attention between Mission Tasks and Ground Avoidance</u>. If all a pilot had to be concerned about was avoiding the ground, he would invariably be successful. But training has to focus on ways of maintaining the number one priority of avoiding the ground while still meeting the tasks requirements dictated by mission and threats. The pilot must divide his attention between multiple tasks. We already know that successful division of attention is critically dependent on repeated practice, practice at integrating ground clearance with mission and threat requirements. But practice alone is not enough. The low altitude environment imposes some special rules governing how to divide attention.

(1) From a knowledge of the aerodynamics, and the sources of available information, it is possible to determine, for every maneuver at a given AGL, how much free time the pilot has that he can take away from attention to ground clearance without increasing his chances of a collision with the ground. Thus, in a level 6 G turn at 100 AGL, free time is one second, whereas when flying straight and level at 100 AGL, 4 seconds is available as free time. This means that during a turn the pilot cannot focus his attention on any other task than ground clearance for more than one second--he cannot make a radio frequency change, or a computer update, or arm missles, or search the skies for his wingman, or anything else that takes longer than one second at a time. All of those longer tasks must be delayed until after the turn is completed, unless the pilot is willing to gain AGL. Flying in the low altitude environment should be structured by free time limitations, and by a knowledge of how long every task normally takes. (2) A pilot must learn how to husband his time,

especially the time it takes to complete particular tasks. If he fails to complete a task, he cannot persist, but must return to ground clearance first, and then start over on the incomplete task. Further, he must learn where snap glances are sufficient, in lieu of longer extended looks. (3) A pilot has to know when to trade AGL for increased task demands. If free time is not sufficient to perform the tasks demanded by the mission, then altitude must be increased. If threats are pressing, then the pilot can come down, but only if other free time tasks are dropped at the same time. There is a continual interplay between tolerances regarding ground clearance and free time to be used for all other tasks. If a pilot understands this, and is trained to go back and forth, then he can always maximize his free time use while avoiding the ground. (4) As a final and very specific example, a pilot must monitor ground clearance during low altitude turns, and cannot use that time for extended looking at or for wingmen, aggressors, targets, missles they just released, or the beauties of the day. Accident analyses pinpoint turning while looking as the single most likely cause of collisions with the ground.

Summary. I have provided a brief overview of four components of an analysis of low altitude flying. To survive in this environment, while accomplishing his mission, a pilot must know limitations imposed by the aerodynamic consequences of each maneuver made, so he can predict where he will be located at each point in time as a result of the control inputs he makes. Further, he must understand about the kinds of visual and other information available that can inform him about his ground clearance and movement over the terrain, including when that information may be compromised or even misleading. From an understanding of the physics of flight and of the sources of information, a pilot can be trained to extract and process the proper information about their ground clearance in the most efficient manner. Finally, he then can be taught the tradeoff between meeting ground clearance demands and the demands of the mission and threats, so that he has explicit strategies for handling multiple tasks. Some of these four components are well understood, and can be described precisely. Others are presented with less confidence, in the hope that even this tentative description will help focus the needed research to make them precise as well.

Acknowledgment. I gratefully acknowledge the many discussions with the instructors of the 162nd Fighter Weapons School of the Arizona Air National Guard, including Lt. Col. Robert Cassaro and especially with Capt. Milt Miller. Most of the concepts and proposals presented here came from their work, and many of them are already impletmented in their low altitude training program.

References

Miller, M. Low altitude training: How low can you go?. 1983. Tucson: 162nd Fighter Weapons School, Arizona Air National Guard.

PANEL SESSION

PANEL SESSION: FAMILY SUPPORT CENTERS AND THEIR RELATION TO QUALITY FORCE, MISSION SUPPORT, AND READINESS

SESSION CHAIR: J. Robert Stevens (Moody AFB)

PARTICIPANTS: J. Robert Stevens (Moody AFB)

James R. David (Army Community Services Division)

Meg Falk (Navy Family Support Program)

Mel Robertson (USMC Family Program Office)

Rosemary Pezzuto (USCG Family Program Office

Peter H. Neidig and Dale H. Friedman (Behavioral Sciences Associates)

PROCEEDINGS ENTRIES

"Supporting the USAF Mission" (J. Robert Stevens)

"Family matters - everyone is an expert" (James R. David)

"Navy family support program" (Meg Falk)

"Marine Corps family services center" (Mel Robertson)

"United States Coast Guard family programs" (Rosemary Pezutto)

"A civilian contractor's observations of the military environment" (Peter H. Neidig and Dale H. Friedman)

US Air Force Presentation - 9th DoD Psychology Symposium

Presentation made by: Dr. J. Robert Stevens
Director, Family Support Center
347 CSG/FS
Moody AFB, GA 31699
AV 460-3333

TOPIC - Supporting the USAF Mission

After more than two and a half years in operation it has become evident that the Family Support Center can fit into the base structure and become a relevant part of the support mission. This is possible by having the following philosophy: immediacy, problem(s) clarification, prioritization and decision(s), development of an action plan, problem resolution, follow-up and integration with the base units. This enables the Family Support Centers to avoid the treatment approach and concentrate on short term intervention, information and proper referral to existing base and off-base agencies.

In order to be able to work in this environment it is necessary to understand the cohort group of our enlisted ranks between the ages of 18 through 23: young people who reached maturity at the end of the 1970´s. For many of them their educational skills in the basics is troubled to say the least. In relation to the earlier cohort groups of the late 60´s and 70´s I find that these folks are more down to earth and less "messed up." Because we let them come in now with dependents we find that a larger portion of this cohort group is married and has children as opposed to their civilian counterparts. The pressures of being young parents has caused many of them to look at the Air Force as an "eight to four" job.

Unlike their civilian college counterparts they will not have a prolonged adolescence. Although they take on more responsibility than their college cohorts they still exhibit certain adolescent behaviors. The economy and a limited career market has pushed them into being more serious about their future. I think we can do a lot with them as they move through the Air Force if we can assist them in coping with the very hea... responsibilities they have taken upon themselves. A special caveat is necessary at this point. I call it the "three D´s"; which means Don´t Develop Dependency. What we are striving for is to develop "warriors, not weepers", and Family Support Centers can be very instrumental in this process. By getting these young people to improve their skills in problem solving and coping, we can make them more productive and assure that they will have a full career in the Air Force without interruptions.

A common question that comes up is, are Family Support Centers doing the jobs of commanders and first sergeants? They are assisting commanders and first sergeants and enabling them to spend more quality time with their good troops. One of the sad things we notice is that commanders recognize problem individuals whenever they visit a shop; what they don't do is give enough recognition to the solid troops who perform for the unit on an every day basis. If we can assist commanders and first sergeants in caring for those who have problems and getting them back on track, plus identifying those that won't ever make it, commanders can then work procedures for channeling non-productive members out of the service. We have then become a part of the quality control mechanism by freeing commanders and first sergeants to devote their time and energies to those members we want to keep in.

Single Parents - The Non-Issue Issue. Ever since I started in Family Support Centers I was given a lot of materials about single parents. I find that most single parents realize that the military is one of the best places they can be. We don't have to spend large amounts of money to know what single parents would like. Their needs are quite simple; free babysitting and the freedom to take care of their youngsters when emergencies arise. None of the latter are things that the military can provide for single parents. We, just like civilian industry, are unable to create special exceptions in the workplace. Being honest and explaining to single parents that this is part of the hardship they must learn to bear and that we will try to support them as much as we can; then we are assisting them in better coping with the responsibilities they have. Responsibility puts a lot of maturity on single parents. Programs for them are very nice, but more often than not only brings up the problems that they already realize they have. Probably the best thing to do for them is to deal with their needs on a case by case basis and see if they can keep a military career and at the same time be a responsible parent.

Why do we put so many resources into family support. Whether we like it or not more and more of our members are heads of families. We cannot order them to forget those family needs and obligations. Our assets of military hardware are extensive and expensive. Unless we want the biggest static display of hardware the world has ever seen, we need to provide high-touch for our high-tech environment. The support we give has the payoff in better mission readiness and effectiveness.

US Army Presentation - 9th DoD Psychology Symposium

Presentation made by: LTC James R. David, Ph.D.
Chief, Army Community Service Division
Office of the Adjutant General
Headquarters, Department of the Army
Alexandria VA 22331
(202) 325-9390 AV 221-9392

Family Matters - Everyone is an Expert

Theoretical underpinnings of Family Support Center model will be presented and evaluated for continuing relevance in view of societal changes and current Army emphasis. Operational pitfalls such as, control vs. facilitation; my needs vs. commander's desires; mechanic vs. engineer; and replacing vs. complementing unit leaders will be examined. Non-negotiable essentiality of professional staff involvement will be stressed; not everyone is an expert.

I. INTRODUCTION

a. I could tell you about Army Community Services but I won't (because it would be boring).

b. If I was going to tell you about it I would say:

1. Started 1965 to support families of deploying troops

2. 162 centers

3. Roughly 5,000 paid and 5,000 non-paid staff

4. Annual budget of 22 million dollars

5. Decentralized management

6. 892,000 volunteer hours in 1982

7. Managed 4,000 child abuse cases

8. Over 5,000 debt liquidations ($22 million total indebtness)

9. 154,000 persons used Lending Closet

10. 330,000 persons received orientation briefings

11. 840 Foster Care placements

12. Average cost per center 67 thousand dollars (cheap)

c. Instead I want to talk about 3 areas:

1. Theoretical underpinnings of military Family Support Centers

2. Some operational pitfalls

3. Relevance to military operations

II. Discuss

a. Theoretical Underpinnings (Great Society and Employee Assistance Program)

1. Rooted in community

2. Continuing needs assessment

3. Outreach efforts

4. Coordination of like activities

5. Volunteer involvement

6. Serving commanders

b. Operational Pitfalls

1. Mechanic vs. Engineer (sending a boy/girl to do a man/woman job)

2. My needs vs. commander's desires

3. Control vs. facilitation/coordination

4. Replacing vs. supplementing/complementing unit leaders

c. Relevance to Military Operations

1. Army Family White Paper and action plan

2. Sinai Deployment

3. New Manning System

4. Mobilization

5. Research - Walter Reed Army Institute of Research and Army Community Service Program Evaluation

III. Closing

a. Non-negotiable Essentiality of Professional Staff

b. (Title of talk) most do think they are family experts

c. While everyone is a family expert, we must listen, receive their input and evaluate it if we are to succeed.

US Navy Presentation - 9th DoD Psychology Symposium

Presentation made by: Ms Meg Falk
Navy Family Support Program
NMPC-66-OP-156
Navy Annex
Washington DC 20380
AV 224-1006

Topic - Navy Family Support Program

1. Mission - The Navy Family Support Program was established with one fundamental responsibility: Support of the Navy Mission.

The specific mandates for the Navy Family Support Program are to assist commands in improving combat readiness, on the job performance and retention of qualified Navy men and women. This is accomplished by increasing the Navy's awareness of and access to reliable information, resources and services that support and enrich the lives of Navy families and single service members.

2. Program Responsibilities:

a. Family Service Centers

b. Family Advocacy Program

c. Ombudsman Program

d. Inter-cultural Relations

3. Specific Program Examples:

a. Family Service Centers

1. Pre, during and post deployment briefings

2. Assistance in home port changes

3. Return and reunion teams

b. Family Advocacy Program

1. Welcome Baby Program

c. Ombudsman Program

1. The communication link between commanding officers and family members

d. Inter-Cultural Relations

1. Orientation programs for Navy members and their families to overseas living.

US Marine Corps Presentation - 9th DoD Psychology Symposium

Presentation made by: Major Mel Robertson
Family Program Office
Code MPH
Hq US Marine Corps
Washington DC 20380

TOPIC - Marine Corps Family Services Center

Mission - To support unit commanders by providing information, assistance and guidance to military families.

Functions -

Information/Referral

Personal counseling

Financial counseling

Relocation assistance

Pre/post deployment assistance

Family Enrichment programs

Others

Demographic Data on Usage

Camp Le Jeune Experience

Programs as result of 23 October Beirut bombing

Thanatos Group (for families of those killed)

Coordination of visits of family members of casualties to Camp Le Jeune

Spring/summer of 1984 a psychologist (ex-Marine) will be working with

Thanatos Group

Casualty assistance officers

Provide briefings at officers call

US Coast Guard Presentation - 9th DoD Psychology Symposium

Presentation made by: Lt. Rosemary Pezzuto
Family Program Office
Commandant (GPS-2)
United States Coast Guard Headquarters
Washington DC 20539
(202) 426-2665

TOPIC - United States Coast Guard Family Programs

Coast Guard Family Programs have been in operation from April of 1983. In this short period the Coast Guard has developed a three-phase approach to assisting families. The phases are:

(1) Proactive - Education enrichment for families in assisting them to deal with the stresses of military life. These stresses include mobilization, for reservists, operational readiness exercises, deployments and mobility for active duty Coast Guard personnel going from one station to another.

(2) Assessments and Referral - Increasing the awareness of the commands to the problems of families, the ability to assess them, and the means by which to refer them to community agencies for treatment.

(3) Education - Coast Guard family education programs include parenting skills, family wellness, spouse employment development programs, teen workshops, and volunteer development.

The Coast Guard has been collecting as much data as possible on the impact of their programs during this first year. Statistics are available from April 1983 on child abuse and neglect.

One of the central problems that has to be addressed is the fact that most Coast Guard stations are very small. They are isolated from the main Coast Guard Command Centers and have to be very reliant upon community resources. As a start it appeared quite difficult for the Coast Guard with their small stations, to be able to develop what is needed. The Coast Guard has met this by educating commands and local commanders on how to make appropriate referrals to available community agencies. Headquarters. Coast Guard has also sent out a lot of materials to station commanders so that they can act as a small self-contained family services program specifically tailored for their local needs. By monitoring the program and getting feedback from the field, Headquarters, Coast Guard has been able to shape them so that they have maximum effect in serving the special needs of Coast Guard families.

Behavioral Science Associates Presentation - 9th DoD Psychology Symposium

Presentation made by: Dr. Peter H. Neidig, Ph.D.
Dale H. Friedman, M.S.
125 South Ribaut Road
Beaufort, South Carolina 29902
(803) 524-3378

TOPIC - A Civilian Contractor's Observations of the Military Environment

Although it may seem presumptions for civilians to make observations about the military, one is reminded of the fact that although we don't know who discovered water, we can be pretty sure it wasn't a fish.

GENERAL OBSERVATIONS

The military is mission oriented. Individuals are valued for their contribution to accomplishing the task and no one is considered to be indispensable. The cost of rehabilitation must always be weighed against the option of replacement. This is reality and beyond debate. Remember who you are working for and why they have employed you.

One who sets out to change the military system would be well advised to recall G.B. Shaw's thoughts concerning teaching a pig to whistle -- it requires a great deal of effort and only serves to annoy the pig.

It is easier to ride the horse the direction he is going in and it is easier to work with the military system than against it.

Mission readiness, morale and retention are a far more effective basis for discussion than humanitarian appeals. The term, "rehabilitation" is preferable to "therapy" and "training" is more acceptable still.

Communications should be brief, concise and free of psychological jargon. "I-thou encounters" and "dialogues" are not held in the same high regard as they are in civilian social service agencies.

Facts and figures rather than theories and generalities are the coin of the realm.

Promptness and brevity are virtues which although relatively rare in the social services, are highly valued in the military.

PUT IT IN WRITING - It is reasonable to assume that any proposals or other written documents will be widely distributed and carefully scrutinized.

Choosing the right "port of entry" is critical. It is important to get to the levels of authority where a decision can be made but it is important to keep in mind that a project can be sabotaged at almost any level. A solid working relationship requires considerable investment of time, but it is worth the effort.

Physical appearances can be important. First impressions colored by obesity or sloth are difficult.

The military is still a male dominated institution. Everything will be somewhat more difficult but not impossible for a female to accomplish.

The military tends to be somewhat wary of civilians and visa-versa. They may have concerns that one is out to "rip them off" or to interfere in some manner with the primary mission; and there may be an initial reluctance to let one get a peek at the "dirty laundry". This period of wariness is usually short lived.

The chain-of-command and the command's need to know should be considered inviolable. It is better to err in the direction of providing the command with too much information rather than too little.

The military is likely to be far more sophisticated about program evaluation issues than the civilian consumer. Measures of consumer satisfaction are unlikely to impress.

Things are apt to not move at all and then to move very quickly. Appointments are often scheduled for 8:00 am the next morning rather than 1:00 pm in two weeks which seems to be the convention in the civilian sector. Similarly, when a program is finally approved, there may be the expectation that the delivery of services begin immediately.

SPECIFIC OBSERVATIONS RELATED TO DOMESTIC VIOLENCE PROJECT:

The military has recognized the value of the support available through primary groups. The cohesion of small groups is systematically encouraged through training, insuring that groups remain intact during deployments and in generally encouraging group activities. As a consequence, service members are much less reluctant to meet and discuss personal issues in a group setting than are their civilian counterparts.

There is generally an acceptance that the line between one's personal life and that which the military has a legitimate interest in is nor nearly so clearly drawn. A man's home may well be his castle but not when it belongs to his employer.

Domestic disturbances are much more easily quelled with far less risk to the military policeman than in the civilian sectors where domestic disturbances are among the most risky activities they are called upon to undertake.

Military personnel are used to following orders and complaining about following orders. Complaining should not be mistaken for insubordination or non-compliance but rather as a natural and perhaps even healthy manifestation.

One function a male therapist can fulfill is to model alternative forms of acceptable masculine behavior.

Do not be defensive or attempt to conceal failures from the command. Your expectations in terms of cure rates is likely to be higher than theirs. (A batting average of 350 will get you in the hall of fame). Creditability can often be enhanced by freely acknowledging when you have failed to rehabilitate an individual.

In many ways the military is a healthy, restricted population. One is not likely to encounter chronic alcoholics, psychosis or other forms of severe pathology which are relatively common place in the civilian community.

Political Ideology and the Willingness to Enlist

Captain Robert P. Kane
United States Military Academy

Abstract

This paper will contrast the relative impact of patriotic and apatriotic motives on decisions made by a group of college students to enlist in the armed forces. The decision to contrast differences in political ideology is made to assess the impact of the increasing reliance by the military to focus on levels of pay and other market-linked conditions of work to establish military manpower policy. The central thesis of this study is that this reliance on economic incentives disregards the political motives of potential members. The basic research design of the paper is to ask who, by political ideology, is most inclined to enlist under presently constructed conditions and then to repeat the process after offering more attractive enlistment incentives. The hypothesis that students whose responses indicate that they tend to be ideologically apatriotic will show a greater propensity to enlist is rejected. The principal conclusion is that patriotism does play a role in determining the enlistment decision.

Background

A directional construct is often employed by theorists to explain the military and its relationship with civilian structure and norms. The military, modeled in this way, is placed on a continuum varying between "convergence" and "divergence" with the greater society (Moskos: 1973). Variation can occur along three dimensions: membership—how much does the military replicate the formal bureaucratic organization of civilian institutions; skill differentials—how many military tasks are similar to the tasks required by other organizations; ideology--do military members retain the same values of the civilian society. This paper will concentrate on the last of these dimensions, political ideology.

Janowitz (1971) presents a convergent model of the military in his research. The military, in Janowitz' view, has necessarily taken on many of the characteristics of large bureaucratic non-military institutions. Furthermore, technological advances in this century may result in military confrontations on a global level. For these reasons, the military hierarchy must be prepared to consider the <u>social</u> (in addition to the military) implications of decisions in the advice it gives to civilian authorities. Therefore, it is important for Janowitz that members of the military force reflect the values of (converge with) the larger civilian society, and more specifically, the broadly liberal values of contemporary parliamentary democracies.

Huntington (1957), also writing during this period, saw the military as necessarily divergent with civilian society. He theorized that, given the necessary autonomy to do its job, the military would maintain high standards of professionalism. These standards would ensure that the military would remain responsive to civilian control, while at the same time, allowing for the most efficient means to carry out the military function.

In the last decade, with the creation of an All Volunteer Force (AVF), major changes in the structure of the military have taken place that bear upon the convergence/divergence formulation. The founders of the AVF consciously designed policies with the aim of creating a representative military force. To meet these requirements, an appropriate system of economic incentives was established (Bachman: 1973). The critical assumption underlying these manpower objectives was that recruitment would follow rational man and utilitarian economic incentives. Toward this end, a system of enlistment bonuses and increased recruit pay was implemented. Non-economic factors were not part of the model.

The basic thesis of this study is that patriotic motives were incorrectly disregarded by the architects of the AVF and that enlistment decisions are not based solely on decisions by rational actors acting according to a market model. Accordingly, the following hypothesis is advanced. If the economistic rationale is correct, it will be the rational actor, or apatriotic individual who will respond most to better enlistment incentives. In other words, the first hypothesis is that apatriotic individuals will most reflect marketplace conditions.

Method

Individuals selected for the study represented students enrolled in an Introduction to Sociology undergraduate course at Northwestern University during the 1981-1982 and 1982-1983 academic years. The survey was given in the fall of 1981 and spring and fall quarters of 1982. The total population usable for statistical purposes during this period was 927 students. Due to the relative closeness of the three sample periods, all three surveys have been combined and treated as one for statistical and analytic purposes.

To test the hypothesis, it is necessary to understand the model of individual enlistment behavior that is used. Political ideology is placed on a continuum varying from patriotic to apatriotic. The model then assumes that where the student falls on the patriotic/apatriotic continuum will influence the decision to enlist. A more extensive study that used national service as an indirect measure of patriotism used this same research design (Burk et al: 1982).

Patriotism: the independent variable

A measure of patriotism was developed from the questionnaire items which was a fairly clearcut measure of willingness to serve one's country. The item was: "Would you favor or oppose requiring all young people to give one year of service to the nation, with a choice to serve either in the military or in volunteer civilian work?" Another item was a little more focused, but still tapped the underlying sentiment of service to one's country: "Do you favor or oppose the registration of the names of all young men so that in the event of an emergency the time needed to call up men for a draft would be reduced?"

These two measures—national service and draft registration—of the independent variables were scored on a standard five-point scale. To the

degree individuals fell on the strongly agree end of the continuum they were termed patriotic, to the degree they fell on the strongly disagree end they were termed apatriotic. Considerable analysis of the variables discussed here, and many additional ones, clearly indicated that the "no opinion" was not a separate uninvolved grouping, but truly a midpoint category, indeed a benchmark, on the patriotic/apatriotic continuum. This will be apparent in the presentation of the tabular data. This can be considered one important preliminary finding of the study.

Enlistment options: the dependent variable

To show the enlistment tendencies of students, two aggregate variables were constructed. One measured economic propensity and another combined student responses to both economic and educational incentives; therefore, it was an overall measure of enlistment propensity. As an example, to test the first hypothesis (that patriotic motives are diminished in the economistic models of the AVF), a national service question could be used as the measure of patriotic motivation. Consistent with the hypothesis, a comparison between apatriotic and patriotic groupings should indicate a greater percentage of potential enlistees in the apatriotic grouping.

Analysis of data

The national service measure of patriotism and overall enlistment propensity variable are presented below (Table 1). In the most patriotic category (those who strongly favored national service), fifty-eight percent of the students are potential enlistees (indicating they would enlist on one or more of the options in the combined propensity variable). This was the greatest percentage of potential enlistees of any of the categories. At the midpoint of the continuum, those who answered no opinion, only thirty-eight percent are potential enlistees. Further, there is a basic monotonic decrease in the percentage of potential enlistees along the patriotic/apatriotic continuum. Similar results were obtained in the analysis using draft registration as the measure.

In brief, there is little evidence that the apatriotic individuals show a tendency to enlist at greater rates than the patriotic individuals. In fact, the opposite finding is true, the most patriotic show a propensity to enlist at rates higher than the apatriotic categories.

TABLE 1
DISTRIBUTION OF POTENTIAL ENLISTEES ACROSS
IDEOLOGICAL CATEGORIES USING NATIONAL SERVICE AS THE MEASURE

	COMBINED ENLISTMENT PROPENSITY			
%	POTENTIAL ENLISTEES	WILL NOT ENLIST	TOTAL	(N)
Strongly favor	58	42	100	(156)
Somewhat favor	46	54	100	(313)
No opinion	39	62	100	(45)
Somewhat oppose	33	67	100	(198)
Strongly oppose	29	72	100	(191)
Column total %	(44)	(56)		(903)
CHI Square = 24.9	p = .0001			

To sharpen the focus and to test the hypothesis in a more direct fashion, the same independent variable is used but the dependent variable is changed to measure the effects of strictly monetary enlistment incentives These are arranged in order of increasing cash incentives. This data is presented below (Table 2).

This table indicates several interesting findings. First, under the standard enlistment option, with no bonus, only 1.3 percent of the entire sample would enlist. Under the regimental enlistment option, which offered a large bonus and the same monthly pay as the basic plan, the number of potential enlistees increased only slightly, to just 4.2 percent. However, the final choice given to the students under the economic propensity variable was the ultra-economic plan, which offered a slightly higher bonus and a significantly higher monthly payment. Considering this option, the percentage of potential enlistees increased to 20.7 percent.

When comparing the patriotic/apatriotic continuum, there is a tendency for the patriotic categories to reflect greater response to the better enlistment options, the opposite of what was hypothesized. Also, the percentages of potential enlistees increases as one moves up the continuum to the patriotic categories. It appears that even the purely economic incentives will not transcend patriotic motivation.

TABLE 2
DISTRIBUTION OF POTENTIAL ENLISTEES ACROSS
IDEOLOGICAL CATEGORIES USING NATIONAL SERVICE (N=914)

ECONOMIC ENLISTMENT PROPENSITY

%	STANDARD(a)	REGIMENTAL(b)	ULTRA-ECONOMIC(c)
Strongly favor	4	9	35
Somewhat favor	1	4	28
No opinion	2	13	29
Somewhat oppose	1	3	19
Strongly oppose	1	6	23
Column Total %	(1.3)	(4.2)	(20.7)
CHI Square = 33.2	p< .001		

(a) Three-year Enlistment at $700 monthly pay
(b) Four-year Enlistment at $700 monthly pay, plus $7,500 bonus at time of joining
(c) Four-year Enlistment at $2,000 monthly pay, plus $10,00 bonus at time of joining.

If this first hypothesis is incorrect, that is, we find patriotic individuals most responsive to improved enlistment incentives, then it may be beneficial to consider to what extent patriotic individuals will enlist. If the density of patriots becomes too great in the military, the ideological representativeness may become skewed, and the development of a separate military ethos becomes a feasible outcome (Bachman et al: 1977). A second hypothesis, then would be the converse of the first: the most patriotic individual will respond greatest to better enlistment incentives.

It should be noted that the two hypotheses predict quite different outcomes of military organization. If the first hypothesis is correct, those most likely to enlist are occupationalists, who make rational choices to join the military, and, see the military akin to a job. If the second hypothesis is correct, those most likely to enlist are the most patriotic (in the pro-military sense), who could constitute a force of political ideologues that retain few linkages to civilian society.

Summary and Conclusions

The data summarized in this paper disagree with the widely held belief that youth no longer join the AVF for patriotic reasons. The principal conclusion is that patriotism does play a significant role in determining the composition of the AVF. Consequently, explanations of enlistment decisions based solely on the dominance of the market model of the military are not likely to explain entirely the enlistment decision. Quite clearly, there are not just economic factors, but many reasons that go into making a decision to enlist.

Finally, economistic models of the military may lose sight of the importance of normative and non-economic components in the enlistment process. Indications of this were shown in the data, as students in the patriotic categories consistently displayed greater enlistment propensity than those in patriotic categories.

These findings indicate that patriotic motives clearly crosscut and transcend economic and educational incentives. This is true even among college students at an elite university, as their choices reveal patriotic motives when making potential enlistment decisions.

BIBLIOGRAPHY

Bachman, J.G. (1973). "The draft, military service, and national unity" in David Gotlieb (ed.), Youth and Contemporary Society (Beverly Hills: Sage).
Bachman, J.G., Blair, J.D., and Segal, D.R. (1977). The All-Volunteer Force: A Study of Ideology in the Military (Ann Arbor: The University of Michigan Press).
Burk, J.S. and Farris, J.H. (1982). The Persistence and Importance of Patriotism in the All-Volunteer Force (Columbus: The Battle Memorial Institute).
Huntington, S.P. (1957). The Soldier and the State (New York: Vintage).
Janowitz, M. (1971). The Professional Soldier (with prologue) (New York: Free Press).
Moskos, C.C. (1973). "The emergent military" Pacific Sociological Review, Vol. 16, No. 2, pp. 255-280.

Young Germans between Compulsory Military Service and Conscientious Objection

Heinz-J. Ebenrett

Psychological Service of the German Federal Armed Forces

Abstract

Sample inquiries among young Germans give way to the assumption that the present rate of conscientious objectors will still increase. Most of the potential conscientious objectors originate from educated classes of society and show an emotional affinity to the peace movement. The preparedness of the young people in total to render military service does not seem to be jeopardized; with a further increase of the rate of objectors, however, a sociological imbalance with regard to a just distribution of the defense burdens will be imminent.

Background

Since 1979 civil polling institutes carry out an annual sample inquiry by direction of the Federal Ministry of Defense among those age groups who are up for conscription. In each of those inquiries about 1500 young men between 16 and 18 years of age are questioned in individual interviews on relevant attitudes and positions regarding defense policy. The data enable us to show trends of development and to relate them to political events, military relevant decisions or other changes of the exogeneous social factors.

General trends

The Federal Republic of Germany is the country with the lowest birth rate in the world. Already at the end of this decade the number of young men needed will be lower than the present requirements for conscripts.

In view of this perspective it will be of the utmost importance whether the expected deficiency will not even be aggravated by an increasing rate of conscripts who claim their right of conscientious objection and who vote to render civilian substitute service. The Basic Law of the Federal Republic of Germany provides for such an individual decision for reasons of conscience. It is, however, considered to be an exception and must be by no means lead to a situation in which only specific sociological groups will render military service and others will not.

There are indications for such a danger. Already today such young men are overrepresented among the conscripts who have a relatively low education, who have worked prior to their conscription and who come from rural areas. Young men with a secondary school education, pupils and students from large cities, in many cases even without close religious bindings, increasingly tend to conscientious objection.

The existing problems will even be aggravated by current demographic trends. Thus, for example, the portion of the rural population with close

...

religious bindings and a traditionally good defense motivation continues to decrease. On the other hand the portion of young people with secondary education has more than tripled since 1950. In addition, youth organizations and institutions as traditional conveyors of social values and obligations continue to loose their importance. They have been replaced by informal small groups and idols advocating a superficial individualization.

Attitudes towards the Federal Armed Forces

A relatively high majority of two thirds of the young men still consider the Federal Armed Forces to be "very important" or "important", respectively. Nevertheless this degree of acceptance means an essential step backwards as compared to previous years when each time three quarters of the young people supported that opinion:

Fig. 1: Significance of the Federal Armed Forces (answers in percent)

	1978	79	80	81	82	83
very important	26	32	35	27	21	22
important	50	48	47	48	54	42
less important	15	11	8	13	13	18
unimportant, harmful etc.	9	9	9	11	11	18

The losses of esteem of the Federal Armed Forces are accompanied by a decreasing confidence in the defense capability of the Western Alliance.

Fig. 2: An attack from the East... (answers in percent)

	1981	82	83
can be repelled	41	36	30
not sure	47	47	49
cannot be repelled	12	17	21

This increasing scepsis must be seen in connection with increasing reservations against the employment of nuclear weapons and against the United States as allies. Only every fourth of the young people supports an employment of nuclear weapons in wartime. In addition, only every second thinks that we can depend "completely" (17%) or "predominantly" (33%) on the United States in wartime. The presence of US forces in Germany, however, is not questioned by the majority:

Fig. 3: Who profits by the presence of US forces in the Federal Republic of Germany? (answers in percent)

	1982	1983
both sides	52	50
above all we	22	17
above all the United States	18	22
nobody	9	11

...

Furthermore it is interesting that only a minority (34%) of the young people questioned thinks that the Federal Republic of Germany is not militarily threatened. Out of that minority again only every second (16.5%) deems it right that the Federal Armed Forces should not fight in the case of a military attack. As in the past the by far overwhelming majority of the young people still considers a national military defense necessary - although without nuclear weapons.

It is alarming that the manifestation of morale is not accompanied by political and military technical knowledge. The greater such knowledge of the young people is - i.e. the more relevant questions they were able to answer correctly - the greater are, for instance, their doubts about a successful repulsion of an attack by the alliance.

Voluntary enlistments

In our samples the percentage of young people who consider a voluntary prolonged duty had remained the same during the past years with just under 20%. In 1983, however, it decreased considerably to 16%:

Fig. 4: Personal attitude towards military service (answers in percent)

	1977	78	79	80	81	82	83
Conscientious objection	13	10	10	9	9	9	13
objection or military service	22	15	14	13	16	14	19
military service	43	52	59	58	57	58	51
compulsory military service or voluntary enlistment	12	12	10	11	9	10	8
voluntary enlistment	10	11	7	9	9	9	8

In contrast to the trends of the answers in our samples, however, the actual number of applications received by the recruiting detachments has rather increased, especially for officer careers:

Fig. 5: Actual number of applicants

	1981	82	83
applications for officer careers (without medical service)	6,500	8,800	11,300
voluntary enlistment NCOs/lower ranks	58,000	49,000	59,000

A provable connection between the respective number of applicants and the situation on the labor market shows that many volunteers base their decision on a rational cost-benefit analysis. In former years those applicants only tended towards short-time enlistment; in the meantime, however, an increasing number also apply for long-term duty. Nevertheless it can be said that most applicants for long-term duty are characterized in the first line by a strong affinity to the military, i.e. they have a positive emotional relation to the military. In our sample the item which most clearly reflected that emotional relation was the following question: "Can you imagine yourself collecting weapons?"

...

Conscientious objectors

The present increase of the number of options for conscientious objection (see Fig. 4) must be seen in connection with the discussion on rearmament during the past year. Obviously it is rather based on superficial attitudes analogous to the "spirit of the time" than on real decisions of conscience. This assumption is supported by answers on an additional question according to which only every second option would be retained if the civilian substitute service were made "uncomfortable":

Fig. 6: Firmness of the decision for "conscientious objection" (group specific answers in percent)

	1978	79	80	81	82	83
full-time nursing of patients confined to bed	55	52	53	49	52	45
2 years civilian service	42	37	54	53	57	49
quartering in barracks	59	59	62	54	66	55
large distance from the place of residence	54	54	59	53	62	51

Potential conscientious objectors are not evenly distributed among the population but predominantly come from certain social groups:

Fig. 7: Sociodemographic characteristics and inclination to conscientious objection (data as of 1982)

Percentage of potential conscientious objectors:	15 %	61 %
characteristics:	upper classes of elementary school catholic/protestant professional training	high school diploma ("Abitur") large city father: secondary education

In addition it is a characteristic of potential conscientious objectors that they are not members in a sports club and show an emotional affinity to the peace movement.

Peace movement

According to our data only 11 % of the young people questioned consider themselves "a part of the peace movement". An additional 62 %, however, show sympathy for its aims. Only 19 % show little understanding for the peace movement and only 8 % reject it categorically. The following tendencies of the answers clarify basic attitudes:

...

Fig. 8: Opinions about the peace movement 1983 (answers in percent / N = 1525)

	"correct"	"wrong"	not sure
- is necessary because of the arms race	63.0	20.5	16.5
- is composed of idealists	70.0	11.1	18.9
- represents interests of the Soviet Union	13.6	68.8	17.6
- endangers the friendship with the United States	23.5	59.3	17.2

The data prove that a clear majority of the young people questioned has a positive attitude towards the aims and arguments of the peace movement. In this context the basic attitude is, however, neither defeatist nor pacifistic: a similarly vast majority of young people would be willing to repel an attack on the Federal Republic of Germany by force of arms. Furthermore a threat from the East is still considered relevant. To an increasing extent, however, the large-scale armament of both super powers is experienced as just such a threat.

Summarizing contemplation

Many of the observations and current empirical trends can be attributed to an increased "emotionalization" as the basic psychological variable. That emotionalization of young people occurs together, for example, with indications for an "interiorization" (stable partnerships, youth religions), "sensibilization" (civil rights, environmental protection), and "solidarization" (peace movement, Third World). Furthermore, the following three social phenomena which are generally in contrast to the objectives of defense policy are of special military psychological importance:

1. The discussions on the defense capability of the Alliance, the reliability of the United States, the nuclear threat, etc. are characterized by an "atmospheric increase of negative perceptions".
2. The emotionalization is accompanied by an increased polarization. Willingness to compromise and tolerance are decreasing.
3. The feeling of one's own powerlessness promotes regressive behavior patterns in the private sphere and tendencies towards neutralization or nationalization, respectively, in the political sphere.

Those three phenomena seem to characterize attitudes and behavior of young Germans in a special way. The latest inquiry data, however, which show that obviously the young people no longer have a predominantly pessimistic view of their personal future and that for the first time in years a majority expects an overall economic recovery give hope for a "turn of events".

AD-P003 332

THE IMPACT OF INFORMATION QUALITY ON CAREER DECISION MAKING

John R. Bruni Thomas W. H. Backman
Navy Personnel Research and Development Center

Patricia L. Randolph
Texas Christian University

Abstract

The present study examined the use of thirteen information sources for career planning activities in a sample of 2859 Navy Officers. Based on component analysis, four use dimensions were identified and were related to the attributes of perceived source accuracy, honesty, and availability. The results indicate differences in usage frequency, but with source being positively correlated with source attributes. However, separate source regressions made for use indicate that only accuracy and availability of the information source predict source use. These results generally agree with prior research in the decision making area.

Introduction

The relationship of individual decision making effectiveness and information quality has received considerable attention in the organizational literature (Caldwell & O'Reilly, 1983; Culnan, 1983; O'Reilly, 1982). These studies demonstrated that as the quality of the information availability increases, so does the quality of the decisions made. Other studies have also shown that there is a positive relationship between information use and the level of environmental uncertainty (Blandin & Brown, 1977). These findings would suggest that when faced with a highly uncertain environment, the decision maker should make greater use of the available accurate information sources in the decision process in order to reduce uncertainty (Galbraith, 1973). However, a number of authors have noted that there are factors that may interfere with use of different sources to obtain accurate information. It has been suggested that an information source that is perceived as providing more accurate information is more likely to be used (Caldwell & O'Reilly, 1983; O'Reilly, 1982; Zmud, 1978). However, several studies have found that perceived source accessibility, rather than the expected accuracy of the information source, determined both source selection and use (Blandin & Brown, 1977; O'Reilly, 1982). Additionally, perceived trustworthiness has been found to hve a major influence on source use (Griffin, 1967). These studies suggest that under conditions of high environmental uncertainty, perceptions of several source attributes affect both source selection and use.

An area where this research would seem particularly relevant is that of career decision making. Thus, the primary purpose of this research is to investigate the influence of several perceived attributes of information sources on an individual's decision to use a particular source regarding career decisions. While the earlier literature has recognized the importance of information source use, the bulk of this work has focused on initial career choice. Given the dearth of available literature on mid-career decision making, the present study sought to examine the following issues: (a) what is the underlying

dimensionality of information source use, (b) does evidence exist which supports an order of information source preference, and (c) what are the source attributes which determine their preference? More specifically, and based on previous decision making research and work by Holzback and Morrison (1982), *it was expected that information sources could be grouped into three or four primary dimensions.* Additionally, previous research has found differences among decision making in predispositions to use certain types of sources (Holland, Stead, & Leibrock, 1976). It was therefore expected that there would be differences in use of information sources, with primary preferences for the interpersonal channels. Finally, information source use would be expected to positively correlate with source attributes of perceived availability, accuracy, and honesty. However, it also was hypothesized that source availability would be the best predictor of source use.

Method

Subjects and Procedure.

The sample for the present study consisted of 2859 U.S. Navy commissioned officers. The present sample represented a 37% return rate. Subjects were selected by commission year (i.e., time in the organization) based on a random stratified sampling design. The resulting sample represented 30% of the total surface warfare community. When the sample was compared to the total sample on demographics, no significant differences were found. The average age of responding officers was 33.1 years (SD=5.89); pay grade ranged from O-1 to O-5. The modal level of completed education (68%) was 16 years (baccalaureate degree) with the next largest group of officers, some 26%, having completed a masters degree.

Selected officers were mailed the Officer Career Questionnaire, a return envelope, and cover letter describing the purpose of the research. Subjects were informed that participation was voluntary and assured that their responses would be confidential. Quesionnaires with ten percent or more missing data were eliminated. Data for the present research is part of a larger study investigating career related issues among Navy officers. For a more detailed description of sampling procedures and measures see Morrison and Cook (1983).

Measures.

Measures used in the present research included:

Information Source Use. Based on the work of Morrison and Holzback (1982), thirteen sources of career information used in career decision making were identified. These included CO (commanding officer), XO (executive officer), department head, other senior officers "in my community," senior officers "outside my community," peers, detailers, "Perspective," "URL Officer Career Planning Guidebook," "Commanding Officer's Addendum," "Officer Billet Summary," "Navy Times," and public media. For each of the information sources, subjects responded on a 7-point scale (1=very low ... 7=very high) on how frequently they used each source for career planning information and guidance.

Information Source Attributes. To measure source attributes subjects were asked to assess the availability, accuracy, and honesty of each of the information sources. These attributes were rated on a 7-point scale as they applied to each information source.

Demographic information on education and tenure in the navy was obtained from organizational personnel data files.

Analyses.

Separate principal components analyses were conducted for information source use as well as for the source attributes of availability, accuracy, and honesty. Salient variable component scores (Gorsuch, 1974) were calculated by summing responses to variables with loadings of greater than or equal to .40 on each source and attribute component. Paired t-tests for differences between information source use components were conducted to assess differences in patterns of use. Finally, correlational and regression techniques were then used to relate source use and source attributes.

Results

To assess the dimensionality underlying the relationships among the information sources, a principal components analysis of the 13 sources was conducted. Four components with eigenvalues of greater than or equal to 1.0 were found, which accounted for 60% of the trace. The first four components were then rotated to varimax simple structure; component loadings and communalities are shown in Table 1. To facilitate interpretation, only items with loadings of greater than or equal to .40 are presented. The first component reflected use of detailers in combination with the printed materials of "Perspective," "URL Officer Career Planning Guidebook," "Commanding Officer's Addendum," and "Officer Billet Summary." This component was labelled "Formal Career Sources." The second component reflected use of CO, XO, and department head. This component was labelled "Immediate Superiors." The third component, "Other Professionals," reflects use of senior officers "in my community," senior officers "outside my community," and peers. The fourth component, "Informal Public Sources," reflects use of "Navy Times" and public media. These results support the expected multidimensional structure of information source use and are consistent with previous research in decision making. Reliability estimates (coefficient alpha) for the four components were .73, .74, .58, and .73 respectively.

Separate principal components analyses of the source attributes of source availability, accuracy, and honesty resulted in simple structures. In other words, for each attribute four components were identified with similar patterns of item loadings as observed with source use. Reliability estimates for the source attribute components ranged between .58 and .81.

Correlations of education and tenure in the navy with information source use dimensions ranged between -.05 and .01, which indicated that source use was not affected by those variables.

Table 2 presents the means, standard deviations, and correlations among the four use components. Two items are worth noting. First, the intercorrelations were all significant, positive, and of moderate magnitude. Further, these values suggest only limited amounts of "halo" among the sources. Second, an inspection of the means revealed that the primary source of information was Immediate Superiors, followed by Other Professionals, Formal Career Sources, and lastly by Informal Public Sources. When t-tests were used to examine the reliability of differences between pairs of the use sources, all differences were found to be significant ($p < .01$). These results provide clear support for the expectation that individuals would differ in their preference for source, with greater use of interpersonal sources (Immediate Superiors and Other Professionals).

Relationships between the four information use dimensions and dimensions of the source attributes are summarized in Table 3. All the correlations were significant and were in the expected directions. High source use was associated with perceptions of greater accuracy, honesty, and availability. However, except for Immediate Superiors, correlations of source use with source accuracy were larger than with availability, which was counter to expectations.

In order to further investigate the use-attribute relationship each information source use dimension was regressed on its respective three attributes. All four regression equations are significant. For three of the four equations the regression weight for honesty was non-significant; for the dimension of Immediate Superiors, significant but small. Perhaps the most noteworthy finding was that the regression weights for both accuracy and availability were approximately equal for two of the use dimensions, Formal Career Sources and Immediate Superiors. Additionally, for the remaining source dimensions, Other Professionals and Informal Public Sources, accuracy rather than availability had the larger weight. This last finding was opposite to that expected.

Discussion

The present study addressed questions of information source use dimensionality and the effects of perceived source attributes on their use for career planning activities. The results of this study supported a multidimensional conceptualization of information source use. Additionally, the source use dimensions were similar to those dimensions that have been found in the decision making literature (Culnan, 1983). Also consistent with previous research, the frequency of use of the four source dimensions indicated a preference for the interpersonal channels of career information. Although not reported here, when use of the individual sources making up the dimension of Immediate Superiors and Oth-

er Professionals were compared, the two persons most interpersonally proximal to the individual were the most used. Of some note was the relatively moderate use of the formal career sources in an individual's career planning activities. This is particularly surprising given that included in this dimension was the detailer, the individual who is the principal organizational representative tasked with assisting the officer in managing his/her career.

The finding that accuracy was either as good or better a predictor of information source use than source availability was unexpected. Before exploring this result, it is important to look at each of the use attribute relationships separately. In terms of the source availability and use relationships, research indicates that this relationship may by reciprocal in nature. Use of an information source improves perceptions of source availability, which in turn leads to increases in its use (Allen, 1977). The importance given to the accuracy of a source is quite understandable given the career implications of decisions based on inaccurate information. Unlike previous decision making research in which individuals making decisions may know that the effects of a wrong decision are often ameliorated by other organizational factors; a bad career decision may be career fatal. O'Reilly (1982) also noted that information in organizations is often contradictory or vague. This is particularly true in terms of career information in organizations where an individual career is subject to periodic changes in organizational career emphasis. Given this about career environments, it appears that more emphasis should be placed on source credibility. A possible explanation can be found in the high correlation between accuracy and honesty.

In summary, the findings of the present study suggest the value of a decision making perspective for studying career decision making. Additionally, the integration of decision making and career literature suggests new areas for research. This latter point can be seen in the present research. Little research has explored the importance of and processes by which an individual manages his/her career after the initial career choice. While the results of the present study are generally consistent with the decision literature, the equal importance of source accuracy and availability is not. Whether this reflects generalizable differences in career decision making versus managerial decision making, or military versus civilian organizational differences could not be answered in this research, and future investigations on these topics are suggested.

References

Allen, T.J. (1977). Managing the flow of technology. Mass.: MIT Press.

Blandin, J.S., & Brown, W.B. (1977). Uncertainty and management's search for information. IEEE Transactions on Engineering Management, EM-24, 114-119.

Caldwell, D., & O'Reilly, C. (1974). The impact of accurate information on job choice and turnover decisions. Proceedings of the American Academy of Management, Dallas, TX.

Culnan, M.J. (1983). Environmental scanning: The effects of task complexity and source accessibility on information gathering behavior. Decision Science, 14, 194-206.

Galbraith, J.R. (1973). Designing complex organizations. Mass: Addison-Wesley.

Griffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. Psychological Bulletin, 68, 104-120.

Gorsuch, R.L. (1974). Factor analysis. PA: Saunders.

Holland, W.E., Stead, B.A., & Leibrock, R.C. (1976). Information channel/source selection as a correlate of technical uncertainty in a research and development organization. IEEE Transactions of Engineering Management, EM-23, 163-167.

Holzback, R.L., & Morrison, R. (1979). Surface warfare junior officer retention: The assignment process. (NPRDC Tech. Rep. 80-13) San Diego, California, Navy Personnel Research and Development Center.

Morrison, R.F., & Cook, T.M. (1983, August). Military officer career development and decision making: A multiple-cohort longitudinal analysis of the first twenty-four years. Paper presented at Paper presented at the Annual meeting of the American Psychological Association, Anaheim, CA.

O'Reilly, C. (1982). Variations in decision makers' use of information sources: The impact of quality and accessibility of information. Academy of Management Journal, 25, 756-771.

Zmud, R. (1978). An empirical investigation of the dimensionality of the concept of information. Decision Science, 9, 187-195.

Note. The opinions expressed in this article are those of the authors, are not official, and do not reflect the views of the Navy Department.

TABLE 1

Principle Component Analysis for Information Source Use

Items.	Rotated Component Structure I	II	III	IV	h^2
1. CO		.85			.75
2. XO		.85			.75
3. Department Head		.64			.53
4. Other Senior Officers in my Community			.74		.64
5. Senior Officers outside my Community			.72		.56
6. Peers			.64		.45
7. Detailers	.49				.44
8. "Perspective"	.72				.54
9. "URL Officer Career Planning Guidebook"	.79				.63
10. "Commanding Officer's Addendum"	.71				.57
11. "Officer Billet Summary"	.63				.43
12. Navy Times				.78	.71
13. Public Media				.85	.77

Note: Only those items with loadings ≥ .40 are reported.

TABLE 2

Means, Standard Deviations and Correlations

Information Source Use	$\bar{X}$	SD	1	2	3	4
1. Formal Career Sources	3.28	1.26	-			
2. Immediate Superiors	3.97	1.77	.25**	-		
3. Other Professions	3.81	1.45	.32**	.38**	-	
4. Informal Public Sources	2.85	1.73	.34**	.10**	.21**	-

**$p < .01$.

TABLE 3

Relationship Between Information Source Use and Source Attributes

Source Attributes	Information Source Use: Formal Career Sources r	Formal Career Sources β[a]	Immediate Superiors r	Immediate Superiors β	Other Profes-sionals r	Other Profes-sionals β	Informal Public Sources r	Informal Public Sources β
Accuracy	.40	.27**	.49	.36**	.50	.39**	.55	.54**
Honesty	.37	.01	.38	-.06*	.34	-.04	.45	-.02
Availability	.39	.24**	.51	.38**	.45	.27**	.26	.11**
R =		.45		.58		.55		.56
R^2 =		.21		.34		.30		.32

Note. All zero order and multiple correlations were significant at $p < .01$

a. Estimated values are standardized regression weights.

* $p < .05$.

** $p < .01$.

The Stability of Vocational Interest - Career Examination (VOICE) Scores

Michael D. Matthews
Department of Behavioral Sciences
United States Air Force Academy

Thomas W. Watson
Air Force Human Resources Laboratory
Brooks AFB, Texas

ABSTRACT

The Vocational Interest - Career Examination (VOICE) is designed to assess vocational interests and to predict job satisfaction in various career fields for Air Force enlisted accessions. An issue that has practical and applied ramifications concerns the stability of interest scores over an extended period of time. Previous research on commercially available interest inventories indicates relatively high test-retest correlations over periods of time ranging up to 35 years. However, these inventories were normed on and designed for so called "white collar" occupations, and stability analyses utilized college student or college trained populations. The VOICE, in contrast, is designed to assess vocational interests for "blue-collar" jobs representative of work activities in enlisted occupations. Accordingly, the purpose of the current study was to assess the stability of VOICE scores over a two to three year period. A sample of 932 1979 and 1980 enlistees were administered the VOICE on the sixth day of Basic Military Training (T_1) and again two to three years later (T_2). Overall, the results indicated moderate T_1 - T_2 correlations, ranging in value from .52 to .70. Preliminary analyses also indicated that the scores of females were somewhat less stable than those of males, especially for nontraditional career/interest areas (e.g., heavy construction, mechanics). The results have implications for the operational use of the VOICE, especially as a counselling tool for personnel changing career fields following the completion of their initial tour of duty. Additionally, the results have theoretical implications concerning the malleability of vocational interests as a function of experience.

INTERELATIONS AMONG CAREER COUNSELING, ORGANIZATIONAL COMMITMENT AND CAREER SATISFACTION

Thomas W. H. Backman and John R. Bruni
Navy Personnel Research and Development Center

Patricia L. Randolph
Texas Christian University

ABSTRACT

The present research is an exploration of the relationships among career counseling, organizational commitment, and career satisfaction. Questionnaires returned by 5051 Navy Aviation Officers contained items forming the above three scales as part of a larger study of career retention of high quality Navy Officers.

Results indicate that there is a strong correlation between organizational commitment and career satisfaction, but the relationships of these two correlates with career counseling is weak. Also, an argument is made that career counseling is actually career advice confounded with internal recruiting. Implications for a redefinition of career counseling are discussed.

INTRODUCTION

Career counseling can best be described as a subset of the relatively new concept of career management (Gutteridge, 1976; Morrison & Holzback, 1980; Pinto, 1976; Schein, 1978); and here refers directly to the interactions of an individual with organizational representatives regarding the organization's career system, norms and values, and opportunities. For example, the Navy has created the position of the detailer who along with the officer's CO and other superiors, are charged with counseling officers and directing them on appropriate career pathways. The detailer is actually a career manager, with counseling being one of many functions in the career management system (recruitment, performace assessment, etc. being others) (Morrison and Holzback, 1980). With appropriate career counseling, career satisfaction and organizational commitment have been hypothesized to increase (cf. Aviation Commanding Officers' Fact Book), but to date this relationship has not been empirically tested.

Organizational commitment has been found to have important implications for a number of individual and organizational variables, such as job absenteeism and turnover, attitudes and expectations. Organizational commitment is here defined as "the relative strength of an individual's identification with and involvment in a particular organization," after Porter and Steers (1973) and Mowday, Porter and Steers (1982).

This concept includes the factors of:

a) a strong belief in and acceptance of the organiation's goals and values;

b) a willingness to exert considerable effort on behalf of the organization;and

c) a strong desire to maintain membership in the organizations (Mowday, Porter, & Steers; 1982).

Although a behavioral component is implied (see b above) organizational commitment is, in fact, an attitude, that has been positively related to performance (Porter, Crampon, & Smith, 1976) and tenure (Mowday et al., 1979; Steers, 1977) and negatively related to

absenteeism (Mowday et al., 1979; Steers, 1977) and turnover (Angle & Perry, 1981; Hom, Katerberg, & Hulin, 1979; Koch and Steers, 1978).

Organizational commitment has been of interest to the Navy's career counseling system because of its particular relationships to performance, turnover, and tenure. With the latter variable, however, it is difficult to discern whether tenure enhances commitment or commmitment directly influences tenure. For this reason, the present study proposes to examine the relationships between organizational commitment, career counseling and career satisfaction while partialling out year of service (and/or rank/grade).

Career satisfaction is an individual's emotional reaction to his/her career. It is expected that career counseling will be positively correlated with career satisfaction. Based on the assumption that career counseling should help the individual's clarify their career aspirations and/or assist them in finding ways of achieving their goals within the organization. This form of counseling should lead to a reduction in career ambiguity which in turn would lead to greater satisfaction.

METHOD

Subjects and Procedure

The sample of 5051 U.S. Navy pilots and Navy flying officers were served with a career questionnaire (Morrison & Cook, 1983). The present sample represented a 41% overall return rate, 38% for the pilots and 48% for the flying officers. Subjects were selected by commission year (i.e, time in the organization) based on a stratified random sampling design. The selected officers representing pay/rank grades 0-1 to 0-5 were mailed the *Officer Career Questionnaire*, a return envelope, and cover letter describing the purpose of the research. Subjects were informed that participation was voluntary and assured that their responses would be confidential. Questionnaires with 10% or more missing data were eliminated. Data for the present research is part of a larger study investigating career related issues among Navy officers. For a more detailed description of sampling procedures and measures see Morrison and Cook (1983).

Measures

Organizational Commitment was devloped from the scale by Mowday, Porter, and Steers (1982). Career Satisfaction was developed from the scale by Morrison and Cook (1983).

A measure of the amount of career counseling received was derived from a set of questionnaire items such as "I've been counseled on how the Navy's system works for members of my community" (agree-disagree 7 pt. scale) and "I've had good counsel on the Navy's norms and values for officers" (agree-disagree 7 pt. scale). Only those officers whose grades were between LTJG and CDR were included due to the ensigns' low knowledge of the Navy career system, along with the low sample of respondents in this grade.

Principal component analyses were conducted on the various statements. Items which loaded highest on the factor that best represented the overall construct of the amount of career counseling received were then selected to form a career counseling scale. Items within this scale that were found to have the highest estimate of internal consistancy reliability were then chosen to construct the final composite score.

TABLE 1

Principle Component Analysis (Varimax rotation) of Career Counseling

Items.	Component Structure I	II	h^2
1. Importance of having someone avaliable with whom I am comfortable and trust to discuss my career.		.79	.63
2. Officers need a special career counseling system for them.		.75	.59
3. My senior officers interact with me frequently	.58		.35
4. I've been counseled on how the Navys * career system works for members of my community	.81		.66
5. I've been counseled on the Navy's career * opportunities outside of my community.	.69		.49
6. I've been counseled on the timing and * proper career progression which will help me reach my career goals in the Navy.	.84		.71
7. I've had good counsel on the Navy's norms * and values for officers.	.77		.59
Eigenvalue	2.81	1.22	
Percent of Variance	40	17	(total=57)

* Items used to construct a scale to measure the amount of career counseling received.

Note: Only those items with loadings ≥ .40 are reported.

alpha of component I = .81

alpha of component II = .34

RESULTS

Component Analyses

Utilizing the criteria of valid principal components as those with eigenvalues greater than or equal to 1.0, two emerged from the analysis. The first component (Table 1) was interpreted as measuring the amount of counseling received and the second component as a measure of the perceived importance of career counseling. Of interest to this study is the first component. Five items correlated greater than or equal to .40 with the factor labeled as Career Advice; these loadings ranged from .58 to .86 and accounted for 40% of the variance. However, when an estimate of the internal consistancy was tested for these 5 items, one item was found to slightly decrease the alpha ("my senior officers interact with me frequently"). Therefore, 4 out of 5 items were retained to form the final career counseling scale. The alpha of .81 also represents the lower bound estimate to the Population Reliability based on the formulas of Woodward and Bentler (1978).

Correlation analysis

Correlations among the three scales were calculated across all ranks. Additional analysis were conducted controlling for rank. Rank was found to explain less than 2 percent of the variance, and so is not further described.

TABLE 2

Correlations among scales

variable	n	x	sd	Organizational Commitment	Career Satisfaction	Career Counseling
Commitment	4745	5.0	0.82		.726**	.303**
Satisfaction	4745	5.8	1.0			.311**
Counseling	4745	4.5	1.3			

** $p < .01$.

DISCUSSION

The Amount of Counseling Received scale for the aviation community produced a similiar reliablity and content as measured in the additional officer samples (Backman, unpublished). The scale had a positive and low correlation with the other scales, Career Satisfaction and Organizational Commitment (Table 2). The larger correlation between Organizational Committment and Career Satisfaction might shed some light on causal factors to organizational commitment. Although these analysis were not conducted to show causal relations, an argument can be made that factors which cause a person to become commited to the Navy are also personally satisfying. We examined questionnaire comments from the respondents, interviews from detailers, interviews from CO's and written "counseling materials" to verify and elaborate that to the navy "counseling" is a surrogate word for advice and internal recruiting on how to persuade the individual to select positions that are compatible with the Navy's manning and retention requirements. The term "counseling" is, however, misleading, based on contents of materials provided to the individuals.

Organizational commitment was found to be positively correlated with rank (a surrogote for tenure) but not enough to confirm the findings of Mowday et al. (1979) or Steers (1977). However, the decision to remain in the Navy beyond the minimum service requirement is strongly related to rank (Backman, unpublished data). Thus, organizational committment is dependent upon factors other than tenure. From a research perspective we now ask the

questions: 1) What are the underlying common factors which link organizational commitment and career satisfaction? 2) What are the causal mechanisms of organizational commitment and career satisfaction? Once the factors and mechanisms are understood, we can address the problem of devising a career counseling program, as part of the detailing system, that would aid an officer in choosing the appropiate career path. Such a program must include information on the Navy's retention and manpower requirements, information on methods to increase the individual's organizational committment, and information on methods to increase the individual's career satisfaction. Because of the strong and positive relationship between organizational commitment and career satisfaction it may be possible to target one aspect (for example career satisfaction) and achieve the other objective simultaneously.

REFERENCES

Angle, H., & Perry, J. (1981). An empirical assessment of organizational commitment and organizational effectiveness. Administrative Science Quarterly, 26, 1-14.

Gutteridge, T.G. (1976). Commentary: A comparison of perspectives. In L. Dyer (Ed.), Careers in Organizations. Ithaca, NY: NYSSILR, Cornell University.

Hom, P.W., Katerberg, R., & Hulin, C.L. (1979). Comparative examination of three approaches to the prediction of turnover. Journal of Applied Psychology, 64, 280-290.

Koch,J.L., & Steers, R.M. (1978). Job attachment, satisfaction, and turnover among public employees. Journal of Vocational Behavior, 12, 119-128.

Morrison, R.F., & Cook, T.M. (1983). Military officer career development and decision making: A multiple-cohort longitudinal analysis of the first twenty-four years. Paper presented at the annual meeting of the American Psychological Association, Anaheim, CA.

Morrison,R.F., & Holzback, R.L. (1980). The career manager role. In C.B.Derr (Ed.), Work, family, and the career: New frontiers in theory and research. NY: Praeger.

Mowday, R.T., Porter, L.W., & Steers, R.M. (1982). Employee-organization linkages. NY: Academic Press.

Mowday, R.T., Steers, R.M., & Porter, L.W. (1979). The measurement of organizational commitment Journal of Vocational Behavior, 14, 224-247.

Pinto, P.R. (1976). Commentary: Neglected issues, unanswered questions. In L. Dyer (Ed.), Careers in organizations. Ithaca, NY: NYSSILR, Cornell University.

Porter, L.W., Crampon, W.J., & Smith, F.J.(1976) Organizational commitment and managerial turnover: A longitudinal study. Organizational Behavior and Human Performance, 15, 87-98.

Porter, L.W., & Perry, J.L. (1979, Nov.). Motivation and public management: Concepts, issues, and and research needs. In Prodeedings of the Public Management Research Conference. Washington, D.C.: Brookings Institution.

Porter, L.W., & Steers, R.M. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. Psychological Bulletin, 80, 151-176.

Schein, E.H. (1978). Career dynamics: matching individual and organizational needs. Mass. Addison-Wesley.

Steers, R.M. (1977) Antecedents and outcomes of organizational commitment. Administrative Science 22, 46-56.

Woodward, J.A., & Bentler, P.M. (1978). A statistical lower bound to population reliability. Psychological Bulletin, 85, 1323-1326.

Note. The opinions expressed in this article are those of the authors, are not official, and do not reflect the views of the Navy Department.

AD-P003 334

Identifying Variables Influential to an Officer Promotion Nomination Board

by

Captain Lynn M. Scott

Air Force Manpower and Personnel Center
Randolph AFB, Tx

Abstract

This study surveyed the decision strategy of an Air Force major command promotion nomination board to below-the-zone major. The board members responded to survey questions to provide influence ratings for 15 promotion folder variables. Analyses addressed: time 1/time 2 influence measurements; board members' intraclass correlation; and measuring specific interaction effects using the Policy Specifying modeling technique; and converged on a subset of influential variables. Additional uses of this approach are also discussed.

The Air Force promotion process plays an important role in the officer personnel management system. Promotion results provide valuable feedback to Air Force leaders about the present and future quality of Air Force officers. The promotion process has also been examined emperically. Moore and Trout (1978) developed the "visability theory" of promotion which described how interpersonal networking in addition to performance were necessary for advancement. Glenn (1977) used the judgement analysis technique, JAN, (Christal, 1968) to identify key factors that randomly selected officers incorporated for their "whole person" concept of promotability. Malmstrom et. al. (1982) also surveyed randomly selected officers to identify variables from the Officer Career Brief listings that were important for selection to a command position. These studies have been very valuable in shedding light on the promotion selection process but are limited in specificity. Moore and Trout's theory and its validation are globally oriented explanations of the promotion process, while the results from Malmstrom et. al. are generalized to the promotion process. Finally Glenn's JAN application derived a promotion policy from hypothetical promotion records scored by hypothetical promotion board members.

The purpose of this study was to examine the officer promotion process in greater detail. An Air Force major command (MAJCOM) promotion nomination board was surveyed to determine the effect of 15

promotion folder variables on their nomination decisions. Specific issues examined included: independent importance ratings for each variable; time 1/time 2 measurements of the variables' importance; board members' interrater reliability; an explicit weighting of interactive effects between specific variables.

Method

Subjects

The subjects were 10 Air Force colonels, members of a MAJCOM promotion nomination board. As board members they were assigned to select eligible captains for nomination to promotion below-the-zone to the rank of major. The captains selected by this board would compete with nominees from other MAJCOMs for below-the-zone promotion at a central promotion board conducted at the Air Force Manpower and Personnel Center.

Procedure

The subjects were surveyed continuously during their week-long board activities. The first survey was administered at the beginning of the week to gather independent ratings of influence toward promotion nomination for 15 variables using a 9-point rating scale. The variables, listed in Table 1, were chosen because they were quantifiable, easily identified in the promotion folder and described officers and their careers. A re-administration was conducted at the end of the week for time 1/time 2 measurements. Throughout the week, self-explicated utility ratings (Hoepfl and Huber, 1970) were collected for specific interaction (i.e. multiplicative) effects between the 15 variables. The subjects were instructed to use a 10-point unanchored utility scale to describe "the benefit to the Air Force of selecting a Captain with the listed characteristics for promotion." Interactions were measured by heuristically pairing the variables into three domains: (1) a performance characteristics domain primarily containing variables from the front side of the Officer Effectiveness Report (OER); (2) a total formal education domain containing advanced degree and Air Force training programs completed; (3) a job characteristics domain which describes the characteristics of the officers' assignments. As an example, the performance characteristics domain and four interaction questions are shown in Figure 1. In all, these domains created the necessary framework to measure specific interactions between the variables and constructs using the Policy Specifying technique (Ward, 1977, Ward, Pina, Fast and Roberts, 1979).

Analysis

The analyses addressed four questions about the promotion process. They were: (1) What promotion folder variables initially influenced the board members?; (2) Did the week long experience of reviewing records affect the variables' relative importance ratings?; (3) Were the board members in agreement about the variables' importance at both the beginning and the end of the week?; (4) What variables both independently and interactively enhanced officer's chances for the board's below-the-zone nomination?

Results

The variable "Demonstration of Leadership Ability" had the highest influence rating in surveys at both the beginning and the end of the week. The variables ratings and ranks are listed for each survey administration in Table 1. There were some changes in the variables' rank-order between surveys but a Spearman Rank-Order correlation indicated that a week of reviewing records had very little effect on influence order ($\rho = .76$, $p < .01$). The intraclass correlation coefficients, r_{11} , (Winer, 1971) for the first survey was .94 and .93 for the end of week survey.

The most pronounced interaction was within the performance characteristics domain between the variable "Demonstration of the Ability to Handle Stress" and the construct "Growth Potential".[1] The board's explicit weighting of this interaction revealed the officers will have low promotion potential if either have a low ability to handle stress (regardless of growth potential) or low growth potential (regardless of their ability to handle stress). Conversely, promotion potential was enhanced multiplicatively for increasing levels of both growth potential and an officers demonstrated ability to handle stress.

Discussion

The benefit of this approach over the research described earlier is that it combines elements from the three studies. The surveys' results converge on a subset of data that can be developed into a global description of the type of officer likely to be nominated; incorporates importance ratings from promotion folder data; and measures specific interaction effects which can be used with the policy specifying technique to develop promotability equations. Most important, however, is that the data were gathered from actual board members during their activities as board members.

---

[1] The construct "Growth Potential" is a synthesis of the following variables: Demonstration of Creativity; Demonstration of Written Comm Skills; Demonstration of Oral Comm Skills; Demonstration of Leadership Ability; Demonstration of Management Ability; and Number of Decorations.

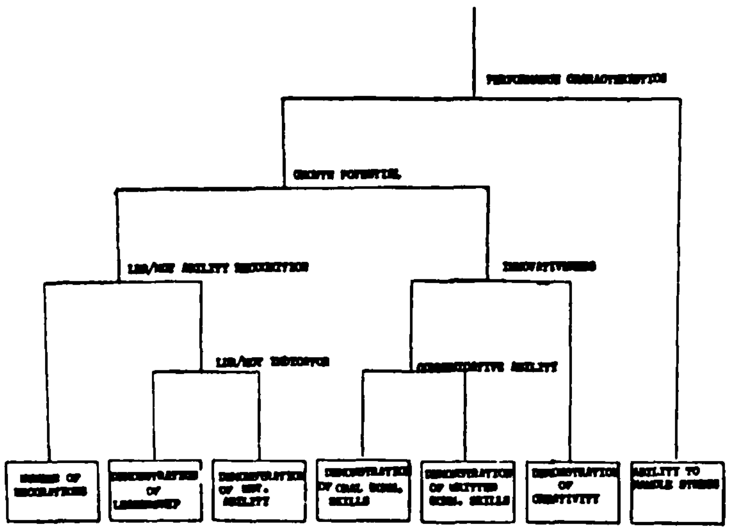
TABLE 1

PROMOTION FOLDER VARIABLES

VARIABLE NAME	Time 1 Measure $\bar{X}$	Time 1 Measure ___	Time 1 Measure Rank	Time 2 Measure $\bar{X}$	Time 2 Measure ___	Time 2 Measure Rank
Demo. of Leadership Ability	7.8	1.03	1	7.6	1.34	1
Demo. of Management Ability	7.1	.998	2	7.3	1.41	2
Assignments in High Level Jobs	6.6	.966	3	6.5	1.17	4
Demonstration of Creativity	6.3	1.33	4	5.6	2.01	10.5
Asgn. in High Level Organizations	6.2	1.31	5	6.0	1.56	6.5
Assignment Diversity	6.1	1.44	6.5	5.6	1.34	10.5
Advanced Degree	6.1	.940	6.5	6.3	.823	5
Demo. of an Ability to Handle Stress	6.0	1.56	8	6.7	1.41	3
Demo. of Oral Communication Skills	5.7	1.05	9	5.8	1.47	8.5
Demo. of Written Communication Skills	5.6	1.17	11	6.0	1.26	6.5
Aeronautical Rating	5.6	.699	11	5.4	1.89	12
Number of Decorations	5.6	1.42	11	5.3	1.41	13
Special Air Force Sponsored Training	5.2	1.22	13	5.8	1.31	8.5
Number of Assignments	4.9	1.66	14	5.2	2.04	14
Overseas Duty	4.8	.942	15	5.5	1.50	15

RATING SCALE: DEGREE OF INFLUENCE

1 = Extremely Low
2 = Very Low
3 = Low
4 = Slightly Low
5 = Neither High nor Low
6 = Slightly High
7 = High
8 = Very High
9 = Extremely High



LEADERSHIP/MANAGEMENT INDICATOR

What is the relative payoff (1-10) for an individual when his record reveals few demonstrations of Leadership and few demonstrations of Management Ability?_____

What is the relative payoff (0-10) for an individual when his record reveals few demonstrations of Leadership and many demonstrations of Management Ability?_____

What is the relative payoff (0-10) for an individual when his record reveals many demonstrations of Leadership and few demonstrations of Management Ability?_____

What is the relative payoff (0-10) for an individual when his record reveals many demonstrations of Leadership and many demonstrations of Management Ability?_____

Figure 1. The Performance Characteristics Domain and an example four interaction questions.

Although the results may not appear surprising to some, they actually provide insight to factors that enhance promotability of officers in this particular command; namely the factors found on the front side of the OER. By collecting measures at all board activities, important career progression variables can be identified and emphasized in each officer's career plan. Assignments can be monitored and particular jobs enriched to provide substantive progression in areas identified from this type of promotion board analysis.

Additional benefits will come if this approach is used Air Force wide. Major Command nomination boards' influential variables and their interactions can be compared with one another as well as with the Air Force central board's influential variables. The total effect will be a more productive force with a career development program more closely aligned with the "whole person" concept that is the corner stone of all selection and promotion boards.

References

Christal, R.E. JAN: A technique for analyzing group judgement. PRL-TDR-63-3, AD-403-813. Lackland AFB, Tx: 6570 Personnel Research Laboratory, Aerospace Medical Division, February, 1963.

Glenn, P.E. A judgement analysis approach to examining a whole person concept of air force promotions. AD-A056499. Wright-Patterson AFB, OH: Air Force Institute of Technology, December, 1977.

Hoepfl, R.T. and Huber, G.P. A study of self-explained utility models. Behavioral Science, 1970, 15, 408-414.

Malmstrom, F.V., Batchelor, E., Reamy, C.J. and Coffman, R.M. Selection for command: what really counts. Paper presented at the Eighth Psychology in the Department of Defense Symposium, USAF Academy, CO. 1982.

Moore, D.W. and Trout, B.T. Military advancement: The visability theory of promotion. American Political Science Review, 1978, 72, no. 2, 452-469.

Ward, J.H., Jr. Creating mathematical models of judgement processes: from policy capturing to policy specifying. AFHRL-TR-77-47, AD-A048983. Brooks AFB, Tx: Occupational and Manpower Research Division, Air Force Human Resources Laboratory, August 1977.

Ward, J.H. Jr., Pina, M., Jr., Fast, J.C., and Roberts, D.K. Policy specifying with application to personnel classification and assignment. Paper presented at the meeting of the Military Testing Association San Diego, October 1979.

Winer, B.J. Statistical principles of experimental design. New York: McGraw-Hill, 1971.

The opinions expressed in this paper are those of the author and are not necessarily the position of the Air Force Manpower and Personnel Center.

THE ROLE OF RETENTION IN MANAGING THE AVIATION WARRANT OFFICER FORCE

Sandra S. Martin
Anacapa Sciences, Inc.
Fort Rucker, Alabama

Michael G. Sanders
U. S. Army Research Institute Field Unit
Fort Rucker, Alabama

ABSTRACT

A survey conducted by the U. S. Army Research Institute (ARI) in Fiscal Year 1979 identified the number and types of aviation warrant officers (AWOs) who were leaving the Army and the factors that influenced the AWOs' decisions to leave. Despite an increase in retention of AWOs since the survey was conducted, the need for continuing concern about AWO retention exists. ARI currently is developing a separation questionnaire that will be administered to all AWOs who leave the Army. Data yielded by the questionnaire, along with selection and classification data, will be used to establish an AWO Force Management System. The system will enable the U. S. Army Military Personnel Center to react more rapidly and more appropriately to deficiencies, as well as overstrengths, in specific occupational specialties. The separation questionnaire development is discussed and a training cost analysis that demonstrates the value of the system is presented.

BACKGROUND

In Fiscal Year (FY) 1979, the U. S. Army Military Personnel Center (MILPERCEN) noted a trend toward decreased retention of Aviation Warrant Officers (AWOs). Retention data indicated that, for those AWOs who had graduated from flight training in FY 1976 and FY 1977T and who were eligible to leave the Army in FY 1979, retention beyond initial obligation was approximately 45 percent. In contrast, during the three previous years, retention of AWOs at the same career point had remained relatively constant at approximately 65 percent (see Figure 1).

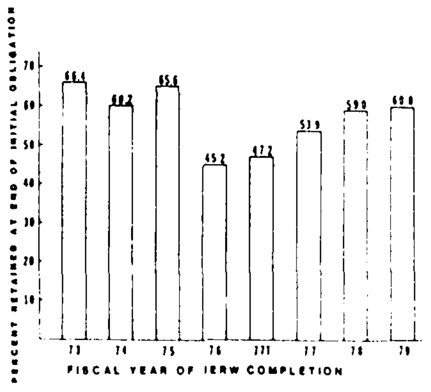


Figure 1. Retention of AWOs at the end of initial obligation.

The research reported here was supported by the U. S. Army Research Institute for the Behavioral and Social Sciences. The views, opinions, and findings are those of the authors and should not be construed as an official Department of the Army position, policy, or decision.

MILPERCEN was concerned that, if the trend toward a decrease in retention of AWOs continued, the Army's aviation readiness and combat effectiveness would be seriously reduced. The concern prompted MILPERCEN to request that the Army Research Institute (ARI) provide research support to investigate AWO attrition. In response to MILPERCEN's request, ARI conducted a worldwide survey of Army aviators that identified (a) the number and types of AWOs who were leaving the Army, and (b) the factors that influenced the AWOs' decisions to leave.

The ten factors that attritees identified as having the most influence on their decisions to leave the Army can be classified into three major areas of concern: (a) pay and benefits (e.g., unequal flight pay, erosion of benefits, etc.), (b) supervision and leadership (e.g., lack of technical knowledge about aviation matters by the chain of command, etc.), and (c) assignment and career factors (e.g., lack of opportunity for assignments to desirable installations, uncertainty about future career opportunities in the Army, etc.). These factors subsequently became the focus of a series of initiatives developed by MILPERCEN to improve retention of AWOs. The most publicized initiative was equalization of flight pay between warrant officer and commissioned officer aviators. The research and the resulting initiatives are described in detail in a series of U. S. Army Aviation Digest articles (Everhart & Sanders, 1981; Morgan & Johnson, 1981; Rogers & King, 1981; Sundy, Ruffner, & Wick, 1981).

IMPACT OF THE INITIATIVES

Since the initiation of the AWO retention research in FY 1979, retention of AWOs has steadily increased. The retention rate for first-term AWOs who were eligible to leave the Army in FY 1980 was approximately 54 percent; the retention rate for first-term AWOs who were eligible to leave the Army in FY 1981 was approximately 59 percent; and the retention rate for AWOs who were eligible to leave the Army in FY 1982 was approximately 60 percent (see Figure 1). Because of the transition from a 3-year to a 4-year initial obligation for AWOs who began flight training after 1 October 1978, few first-term AWOs were eligible to leave the Army in FY 1983.

Feedback from individuals in the field suggests that the continued increase in AWO retention is due, in part, to the Army's concern as expressed in the retention initiatives that were enacted during FY 1982. In addition, there is evidence that the decline in the economy has limited the availability of civilian jobs during recent years. The decreased chances of finding a civilian job might have encouraged retention of AWOs who would have chosen to leave the Army.

REASONS FOR CONTINUING CONCERN

Despite the recent increase in retention rate, there are reasons for continuing concern about AWO retention. One of the primary concerns is the high cost of training each time the retention rate declines. For example, in response to the high AWO separation rate in FY 1979, the Department of the Army (DA) directed the U. S. Army Aviation Center (USAAVNC) to increase the AWO training rate from 420 in FY 1979 to 853 in FY 1983. Figure 2 illustrates the dramatic increase in AWOs trained at USAAVNC over the past few years.

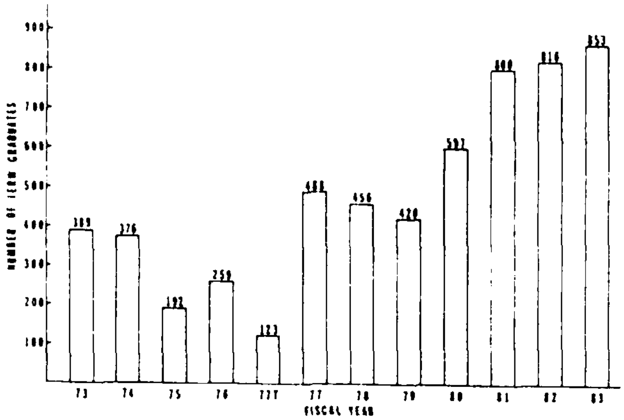


Figure 2. Number of active Army AWO graduates from IERW flight training by fiscal year.

While a higher training rate increases the number of AWOs who remain in the Army, mathematically it also increases the number of AWOs who may leave the Army--even with an improved rate of retention. For example, in FY 1979, when retention rate for first-term AWOs was 45.2 percent, the Army lost 142 AWOs at the end of initial obligation. However, a projected retention rate of 60 percent for the AWOs trained in FY 1980 would represent a loss of 239 AWOs in FY 1984.

The problem of AWO retention becomes of even greater concern when the losses are viewed in terms of training replacement costs. An AWO who separated from the Army at the end of initial obligation in FY 1979 represented a minimum training replacement cost of $189,111 (see Table 1): an AWO trained in FY 1983 will represent a minimum training replacement cost of $254,661 at the end of initial obligation (see Table 2).

TABLE 1

FY 1979 ESTIMATED MINIMUM REPLACEMENT COST OF A UH-1 AVIATION WARRANT OFFICER

COST OF IERW	YEARS OF SERVICE AS AVIATOR (INITIAL OBLIGATION)	FAC 1 ANNUAL FLIGHT HOUR REQUIREMENTS	DOD ESTIMATED COST PER FLIGHT HOUR	ANNUAL COST OF PROFICIENCY TRAINING (COL 3 x COL 4)	CUMULATIVE COST OF PROFICIENCY TRAINING (CUMULATIVE TOTAL OF COL 5)	TOTAL CUMULATIVE COST OF TRAINING (COL 1 + COL 6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
$121,431	1	96	$235	$22,560	$ 22,560	$143,991
121,431	2	96	235	22,560	45,120	166,551
121,431	3	96	235	22,560	67,680	189,111

TABLE 2

FY 1983 ESTIMATED MINIMUM REPLACEMENT COST OF A UH-1 AVIATION WARRANT OFFICER

COST OF IERW	YEARS OF SERVICE AS AVIATOR (INITIAL OBLIGATION)	FAC 1 ANNUAL FLIGHT HOUR REQUIREMENTS	DOD ESTIMATED COST PER FLIGHT HOUR	ANNUAL COST OF PROFICIENCY TRAINING (COL 3 x COL 4)	CUMULATIVE COST OF PROFICIENCY TRAINING (CUMULATIVE TOTAL OF COL 5)	TOTAL CUMULATIVE COST OF TRAINING (COL 1 + COL 6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
127,173	1	96	$332	$31,872	$ 31,872	$159,045
127,173	2	96	332	31,872	63,744	190,917
127,173	3	96	332	31,872	95,616	222,789
127,173	4	96	332	31,872	127,488	254,661

The loss of 142 AWOs in FY 1979 represents a total loss of approximately $27,000,000 (142 AWOs multiplied by $189,111 training cost per aviator; see Table 3). In contrast, a projected loss of 239 aviators in FY 1984 would represent a total loss of approximately $61,000,000 (239 AWOs multiplied by $254,661 training cost per aviator; see Table 3). Since the projections of future losses are not based on inflated dollars and do not include the costs of additional aircraft qualification courses, the actual loss represented by these aviators will be much greater.

TABLE 3

ESTIMATED NUMBER AND TRAINING REPLACEMENT COST OF AWOs WHO LEAVE THE ARMY AT THE END OF INITIAL OBLIGATION

FLIGHT SCHOOL COMPLETION	NUMBER OF AWOs TRAINED	END OF INITIAL OBLIGATION	PERCENT ATTRITION	NUMBER OF AWOs WHO LEAVE (COL 2 x COL 4)	ESTIMATED MINIMUM TRAINING REPLACEMENT COST
FY 1976	259	FY 1979	54.8	142	$ 27,000,000
FY 1977T	123	FY 1979	52.8	65	12,000,000
FY 1977	488	FY 1980	46.1	225	43,000,000
FY 1978	456	FY 1981	41.0	187	35,000,000
FY 1979	420	FY 1982	40.0	168	32,000,000
FY 1980	597	FY 1984	40.0 projected	239	61,000,000
FY 1981	800	FY 1985	40.0 projected	320	81,000,000
FY 1982	816	FY 1986	40.0 projected	326	83,000,000
FY 1983	853	FY 1987	40.0 projected	341	87,000,000

CURRENT AWO RETENTION ACTIONS

The financial loss shown in Table 3 supports the conclusion that the Army needs to continue its AWO retention effort. As a part of the Army's ongoing effort to improve the retention of AWOs, ARI currently is developing a separation questionnaire designed specifically for AWOs. ARI was tasked by MILPERCEN to develop the questionnaire as a follow-on to the retention survey. A preliminary version of the separation questionnaire currently is being field tested.

Once the separation questionnaire becomes operational, it will be administered to all separating AWOs as a part of their general outprocessing from the Army. Information provided by the questionnaire will be used to establish a system that yields continuous feedback about AWO attrition. The system will provide current information about the number and types of AWO losses to the Army and about the factors that influence the AWOs to leave the Army.

Information about the AWO losses can be used in the Aviation Warrant Officer Force Management System (see Figure 3) to determine aviator replacement needs, assess training requirements, and forecast AWO force strength. The information can also be used to determine the optimal number of AWOs that must be retained in order to meet the Army's projected AWO requirements at a minimum training rate. Information about the factors that influence attrition can be used to help personnel managers assess the effect of specific policies and decisions on AWO retention and to develop programs to maximize the retention of high quality aviators.

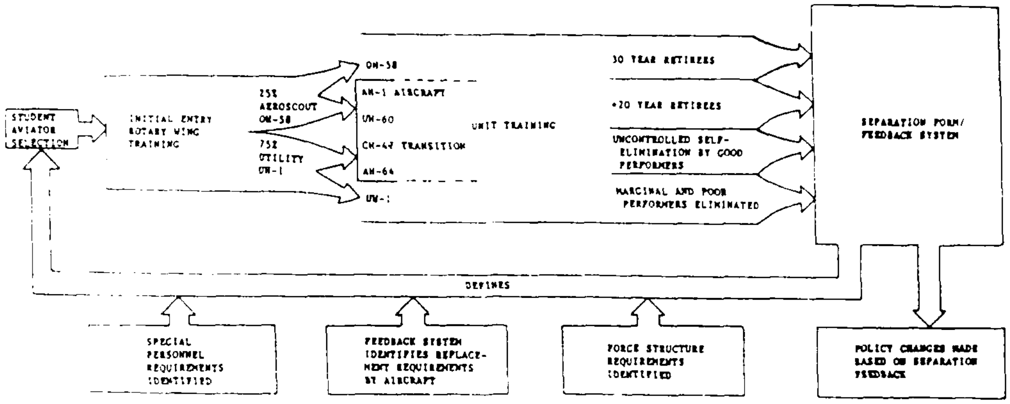


Figure 3. Aviation force management model.

The goal of ARI and MILPERCEN is the development of an Aviation Warrant Officer Force Management System that will have efficient and interactive selection, classification (assignment), and retention programs. Such a system will enable MILPERCEN to react more quickly and more appropriately to specific personnel occupational specialty deficiencies or overstrength situations. The current personnel management system does not effectively address specific occupational specialties or experience levels, nor does it contain a quick feedback mechanism or a data base for long-term personnel projections. Therefore, the AWO understrength problem experienced in FY 1979-81 has become an overstrength problem in 1983. An efficient Aviation Warrant Officer Force Management System will minimize the magnitude and duration of the perturbations experienced above and below the AWO authorization line. The development of such an Aviation Warrant Officer Force Management System is being pursued jointly by ARI and MILPERCEN.

REFERENCES

Everhart, C. D. & Sanders, M. G. (1981). Aviation warrant officer retention: A matter of concern. U. S. Army Aviation Digest, 27(8).

Morgan, G. A. & Johnson, K. M. (1981). Aviation warrant officer retention: A matter of action. U. S. Army Aviation Digest, 27(12).

Rogers, G. L. & King, O. T. (1981). Aviation warrant officer retention: The factors which influence the decision to leave. U. S. Army Aviation Digest, 27(9).

Sundy, B. T., Ruffner, J. W., & Wick, D. T. (1981). Aviation warrant officer retention: An evaluation of demographic items. U. S. Army Aviation Digest, 27(11).

AD-P003 336

The Relationship of Role Congruity to Air Force Pilot Turnover and Performance*

David P. Lohmann
Air Force Management Engineering Agency

Michael J. O'Connell
Air War College

Abstract

The performance and retention of pilot trainees was related to the degree to which their self concept matched the role expectations of others who have performed well in that role. Role congruity was found to be related to performance in pilot training and performance on active duty. The relationship with retention was not significant.

Background

This research explores how and why some new entrants to a new occupation adapt and perform successfully while others do not. We hypothesize that an entrant with a congruous match between his or her own self concept and role expectations of others who have performed well in that occupation will have a greater propensity to stay in that occupation and perform well. A person with a dissonant relationship between self concept and institutional role expectations will be more likely to leave the profession's training institutions or perform poorly in training. If such an entrant does complete training, we expect that he or she will more likely leave the chosen occupation or perform at a lower level than a person with a more congruous match. Others have hypothesized that an individual entering a vocation or an organization is searching for consistency between his or her self concept and the role behavior that the selected vocation or organization demands. According to Super (1957) the young man or woman facing the decision to select a vocation tries to find a medium in which to implement his or her self concept as the person believes it to be and as he or she believes the expectations of the vocation are. Conflict between role expectations and self concept lead to anxiousness and ineffectiveness. As a result, the individual suffering conflict may leave for another career or remain and suffer the frustrations inherent in the continued conflict. Less than optimal performance is one outcome from these decisions.

To explore these relationships the integration of Air Force pilots into a unique and challenging occupation was studied. Two phases of pilot socialization were analyzed: (1) performance and retention in pilot training, and (2) performance and retention during the nine years following pilot training. We hypothesize that completion of and success in pilot training and subsequently in their Air force flying roles is directly related to how closely their self concepts match those of the idealized Air Force fighter pilot. The role of the fighter pilot is suggested as the ideal Air Force pilot behavior model. The keys to success in air combat leadership is captured in the fighter pilot role model. World War II research demonstrated that those most likely to succeed in combat had a love of flying, a desire to contribute, and a sense of aggressiveness which led to a desire to use aircraft to fight (Bond, 1952). This research concluded that the opportunity to satisfy deep aggessive drives was more significant in determining combat success than unit identification, hatred of the enemy, leadership or character dispositions.

*Both authors are on active duty in the U.S. Air Force. The views expressed in this paper are their own and do not necessarily reflect those of the Department of Defense or the U.S. Air Force.

Recent studies of successful NATO pilots found characteristics similar to those of successful World War II pilots, namely, aggressiveness, self-confidence, courage, teamwork, responsibility, determination, and self-discipline (Advisory Group for Aerospace Research and Development, 1973).

Method

Research samples

Three different groups of junior officers entering Air Force pilot training were selected to analyze their adjustment to and performance in this unique training environment: (1) 141 1979 graduates from the Air Force Academy, (2) 161 1980 graduates from the Air Force Academy, and (3) 208 officers from various commissioning sources who were in pilot training in 1972 and 1973. Groups one and two trained at different bases throughout southern United States. Group three trained at Williams Air Force Base in Arizona and was composed of different sections who had started pilot training in different months. When role congruity was collected in 1973 a number of sections were near graduation and had already experienced training losses. Group three was used only to analyze performance in pilot training and turnover and performance during the nine years following completion of pilot training.

Measurement

Role congruity is the degree to which a person's self concept is similar to the role perceptions of the idealized occupational persona in question. It, therefore, requires comparable measurement of both the self concept and the idealized role. The semantic differential was used to measure two concepts "fighter pilot" and "myself." Four bipolar adjective sets were selected, based upon established high factor loadings in previous research, for each of three dimensions in semantic space: evaluative, potency and activity. The twelve bipolar adjective pairs for each dimension were:

Evaluative -- superior-inferior; good-bad; satisfied-dissatisfied; contented-discontented.
Potency -- serious-humerous; strong-weak; violent-gentle; hard-soft.
Activity -- lively-sluggish; active-passive; fast-slow; spirited-lifeless.

Each adjective pair was scored on a seven point scale with "7" indicating the best evaluation, most potent, or most active description. The score for each dimension was the average of the four adjective pairs. Additional details on this instrument and scoring procedure are given in Lohmann (1973).

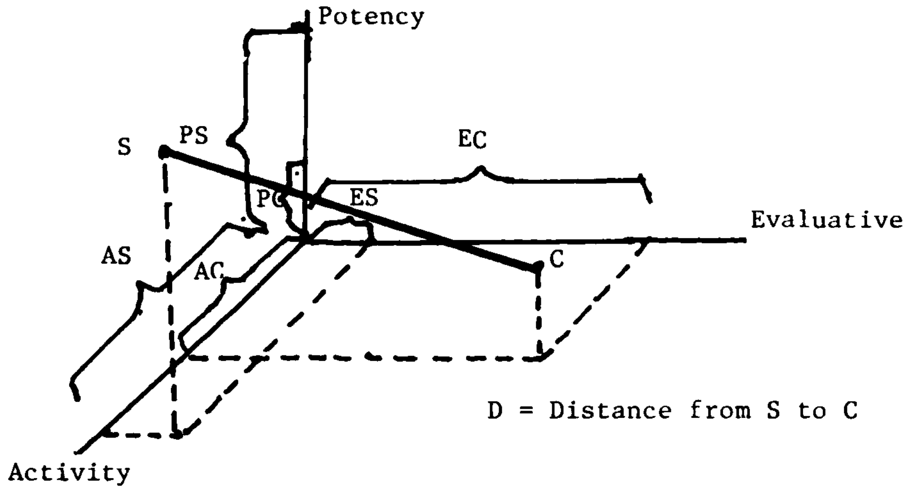
The Euclidian space distance (D) between the individual's self concept and the role standard was used to measure role congruity. An illustration of the D statistic is shown in Figure 1 where:

S = fighter pilot role standard in semantic space
C = individual trainee's self concept in semantic space
ES = evaluative score of the role standard
AS = activity score of the the role standard
PS = potency score of the role standard
EC = evaluative score of the individual trainee's self
AC = activity score of the individual trainee's self
PC = potency score of the individual trainee's self

and --

$$D = \sqrt{(ES-EC)^2 + (AS-AC)^2 + \ldots}$$

Figure 1



Measuring Role Congruity in Semantic Space

The self concept measures, EC, AC, and PC in equation (1), were obtained from samples one and two when these people entered the Air Force Academy in 1975 and 1976, respectively. The same self concept measures were obtained from sample three during their pilot training in 1973. Perceptions of the fighter pilot concept, ES, AS, PS in equation (1), were obtained from 50 operational fighter pilots in 1973. Data for sample three and the fighter pilots were collected in conjunction with research reported in Lohmann (1973).

Turnover. Attrition from pilot training for samples one and two and from active duty for sample three was determined from pilot training and Air Force personnel records. Table 1 summarizes the losses from each group.

Pilot Training Performance. Two performance measures were used: (1) class standing, a composite of academic and flying evaluations expressed as a percentile rank in the class, and (2) average flying grades based on a 4.0 for an excellent checkride, 3.0 for a good one, 2.0 for a fair one, and 0.0 for an unsatisfactory checkride. These data were compiled from records at the Air Training Command Headquarters in San Antonio, Texas.

Table 1

Summary of Losses from Different Samples

Sample	Entered Pilot Training	Lost from Pilot Training	Graduated from Pilot Training	Lost from Active Duty	On Active Duty (Nov 82)
1979 Academy Grad	141	12	129	*	*
1980 Academy Grad	161	22	139	*	*
1973 Pilot Trainees	208**	21	187 (156)***	103	53

* These groups are not allowed to resign from active duty until 1985 for the class of 1979 and 1986 for the class of 1980.

** More than 208 entered pilot training, but at the time the initial data were collected in 1973 only 208 remained.

*** While 187 graduated from pilot training usable data on their Air Force flying assignments were obtained for only 153 pilot training graduates.

Operational Pilot Performance. For sample three a special scale was constructed to reflect the hierarchy of pilot assignments a pilot might eventually achieve. Using records on flying assignments in the nine years following graduation, performance was quantified by assigning points for serving in various positions according to these rules:

Non Rated or Navigator Specialty, Entry Level, Co Pilot	0 points
Aircraft Commander, Instructor Pilot, Flight Examiner, Aggressor Squadron Member, Special Air Mission Pilot	1 point each

Weights were established by having a panel of senior Air Force pilots reach consensus on the relative merits of each flying position. Scores ranged from 0 to 3 points. Scores of two or three indicate that the pilot had successfully passed rigorous screening to identify superior flying performance. Because a pilot might achieve a score of one simply by longevity, we compared 26 high performing pilots scoring 2 or 3 with 127 pilots with scores of zero or one.

Results

Turnover and Performance during Pilot Training

Those who completed pilot training had marginally, but not significantly, greater role congruity with the fighter pilot role standard than those who attritted from training. That the trainee's self concepts were measured four years earlier may explain the weak support for this hypothesis. Because a number of people with self concepts incongruous with the fighter pilot role standard did not enter pilot training, restriction of range may also explain the weak results. Table 2 shows the correlations between the D statistics for the three samples and the two measures of performance in pilot training. The significant negative correlation indicate that greater role congruity is associated with better performance during training.

Table 2

Correlation of Performance in Pilot Training With Role Congruity

Sample	Flying Performance	Class Standing
1979 Air Force Academy Graduates	-.129 *	-.136 *
1980 Air Force Academy Graduates	-.143 **	-.149 **
1973 Trainees	-.119 *	-.132 *

* $p < .10$, ** $p < .05$

Turnover and Performance after Pilot Training

The 103 pilots who resigned between 1973 and 1982 had an average D statistic of 1.76 compared to an average of 1.53 for the 53 still on active duty in 1983 ($p=.12$). This difference shows that stayers had more role congruity in 1973 than did leavers, but the difference only approaches statistical significance . For the 26 who had achieved positions indicating superior ability average role congruity was 1.44, while the 127 who had advanced no further than the position of aircraft commander had an average role congruity of 1.78 ($p<.05$).

Conclusions

This study demonstrates that congruity between a person's self concept and the idealized role standard of a fighter pilot is directly related to successful socialization into and performance in the Air Force pilot occupation. We know from previous research that abilities, prior training, and the nature of flying tasks encountered probably also play a role in the successful transition to a flying career. Additional research is warrented to better understand the dynamics of role congruity overtime, the possible moderating impact of other variables including personality measures, and the ramification of using other role standards or methods of measuring role congruity. Because role congruity is so closely intertwined with the self concept, additional studies should attempt to clarify the differences of these two key variables.

References

Advisory Group for Aerospace Research and Development. "Clinical Psychology and Psychiatry of the Aerospace Operational Environment," Soesterburg, Netherlands, North Atlantic Treaty Organization, 1973.

Bond, Douglas D. The Love and Fear of Flying. New York: International University Press, 1952.

Lohmann, David P. An Examination of Some Behavioral Correlates of Air Force Undergraduates Pilot Training Through the Use of the Porter and Lawler Performance/Satisfaction Model. DBA Dissertation, Arizona State University, 1973.

Super, Donald E. The Psychology of Careers, New York: Harper and Brothers, 1957.

AD-P003 337

PANEL SESSION

Psychophysiological Tools in Engineering Psychology

Emanuel Donchin, Arthur Kramer, Amir Mane,
Demetrios Karis and Earle Heffley

Cognitive Psychophysiology Laboratory
Department of Psychology
University of Illinois
Champaign, Illinois

Introduction

The event related brain potential is a record of the specific response of brain tissue to external, or internal, events. It can be extracted, by computer-aided signal averaging, from the Electroencephalographic recordings that can be obtained from awake, behaving, human subjects. The subject from whom the ERP is recorded can perform, with very little encumberance, fairly complex tasks. Therefore, the ERP provides a unique avenue for observing brain activity that may be associated with cognitive processes. This possibility has become particulalry intriguing when, in the middle 1960s, several "endogenous" components of the ERP were discovered. These components are interesting because they are not obligatory, "sensory", responses to external events. Rather, they seem to be invoked by the processing activities associated with the task the subject is performing.

Within the context of Engineering Psychology these ERPs promise to provide a monitoring tool that is unique in its ability to examine processes activated during task performance even when these processes are not readily observable by more traditional means. However, the language spoken by these components can not be understood without resorting to a detailed analysis of their functional significance. The process involves the conduct of an extensive program, focusing on each component. The first stage of such a program requires the explication of the antecendent conditions for the component. That is, the investigators need to identify the attributes of the task that call for the elicitation of the component. They also need to determine which independent variables control the amplitude, latency and scalp distribution of the component.

From this analysis there emerge various hypotheses regarding the functional significance of the components. These hypotheses suggest the specific information processing activity that is manifested by the component. The theory is tested by testing predictions about the consequences of the component. In the course of these analyses much useful data is acquired that allows the utilization of the ERP as a tool in the study of cognitive function even before the component's significance has been explicated. As will be shown by the presentations in this symposium it is possible to use the amplitude of the P300 component as an index of resource allocation between tasks. It is also possible to use the latency of the P300 to measure the duration of mental activities.

In this introductory talk, we will illustrate the manner in which basic information about ERP components, focusing especially on the P300 component, is acquired. We will endeavor to show how the data accumulated over the last two decades is leading to the view of P300 as a manifestation of updating activities in Working Memory. The manner in which this hypothesis is derived and tested will be discussed. The relation between this approach and the specific application of ERPs in Engineering Psychology will be reviewed.

Selective Attention in Human Engineering

Selective attention is a crucial component of task performance in virtually every man-machine system. Thus, theoretical and practical considerations relating to attention should be a primary focus in several domains within the human engineering of such systems. In the realm of design, a general psychological understanding of attentional capacities and mechanisms should guide the development complex systems. In systems evaluation, limitations in the ability to attend to multiple information sources should be a primary consideration in judging the merits of various prototypes. In the selection of operations personnel for complex systems, individual differences in the ability to selectively attend to relevant information channels can be an important factor in the prediction of future performance. Research on the relation of human event-related brain potentials to selective attention has made contributions to our theoretical understanding of attentional capacities and has yielded several methods for the practical application of ERPs to human engineering problems.

In the domain of theory, the study of ERPs has given insights into the extent to which information is processed in the nervous system, given various stimulus configurations and with differing task requirements. The research has ranged from investigations of "early filtering" of sensory channels (Hillyard and Hansen, 1984) to experiments on higher-level stimulus categorizations in multiple task situations (Kramer, Wickens and Donchin, 1983). Experimental results relating P300 to the processing of unpredictable task-relevant events is particularly applicable to the evaluation of systems which present the operator with multiple sources of information of varying levels of significance. Because the P300 gives an index of information processing that is not dependent upon manual responses, investigators have proposed using P300 to obtain a fine-grained analysis of the depth to which significant signals are processed (Isreal, Wickens, Chesney and Donchin, 1980). Thus, ERPs are beginning to be considered as a tool for the human engineer in the evaluation of complex systems. In the realm of personnel selection, a program of research is investigating the power of a battery of ERP tests to predict future performance of candidates in a task with multiple control and processing elements (Karis, this panel). We are also evaluating ERPs as a means to add further power to attention-switching tests (Gopher, 1982) that when given to fighter pilot candidates have yielded significant correlations between behavioral measures in the task and future success as pilots.

Mental Workload

The P300 component of the ERP has been shown to be a sensitive index of perceptual/cogntive workload. The sensitivity of the P300 to changes in cognitive workload has been demonstrated in both primary and secondary task workload assessment techniques. In the secondary task paradigm, subjects perform a primary task such as monitoring a simulated air traffic control

display for course changes (Isreal, Wickens, Chesney and Donchin, 1980) or tracking an object in three dimensional space (Kramer, Wickens and Donchin, 1983) while also counting covertly the total number of one of two types of visual or auditory events (i.e. bright flash or a high pitched tone) presented in a Bernoulli series. The P300s are elicited by the counted and uncounted secondary task events. Increasing the perceptual or cognitive difficulty of the primary task results in a systematic decrease in the amplitude of the P300s elicited by the secondary task events. The amplitude of the P300 is presumed to index the resources allocated to the secondary task performance and therefore would be expected to decrease with increases in the difficulty of the primary task (for a discussion of resource theory see Kahneman, 1973; Wickens, 1980). If P300 is in fact an index of processing resources then it would be predicted that P300s elicited by discrete primary task events would increase in amplitude with increases in the difficulty of the primary task. The direct relationship between P300 amplitude and primary task difficulty was demonstrated in an experiment in which subjects performed a pursuit step tracking task at different levels of system order (Wickens, Kramer, Vanasse and Donchin, 1983). P300s were elicited by changes in the spatial position of the tracking target.

These results have both applied and theoretical implications. P300 provides a measure of workload which is sensitive to a specific type of difficulty (perceptual/cognitive). In addition the P300 elicited in the primary task does not require an extraneous response and therefore does not intrude on primary task performance. On the theoretical side, P300 has provided support for the models of attention which poist a resource structure underlying overt performance.

Training and Skill Acquisition

The acquisition and performance of a complex skill has been a topic of investigation in the CPL for the last three years. A computer controlled video game "Space Fortress" was developed for research purposes. Briefly, this game requires the player to control a vehicle in a frictionless environment and to use a discrete weapon system to "destroy" enemy objects.

Our first goal was to establish an objective procedure for analyzing a complex task into its components. The additive factors procedure (Sternberg, 1969) was devised to assess the independence of human information processing stages by using reaction time measures. In the present study, we adapted this procedure to determine the interrelations and interactions among components of a complex task and the skills associated with their performance. Specifically, we manipulated orthogonally the difficulty of four aspects of the task, and used analyses of variance to assess the effects of these manipulations on a variety of performance measures. Results indicate that the task can be analyzed into at least three components: appraisal, motor, and perceptual-motor. The appraisal and motor components are essentially isolable. The perceptual-motor component is not isolable since it interacts with the other two components.

This work was followed by an analysis of the susceptibility to stress of the different task components as identified previously. Stress was imposed by requiring our subjects to perform the task for long duration missions, which involved 12 hours of almost continuous performance. Subjects performed both day and night missions. We also compared the performance of both naive and expert subjects. Finally, we incorporated measures of the ERP to determine the utility of these measures as indices of specific psychological resources. These results indicate (a) that some

components of the ERP are consistent over a 12 hour recording period, (b) systematic changes occur in other ERP components over time. There is some indication that these changes are associated with different types of performance decrement.

Another direction of research focused on training methods. Based on the task analysis a set of subtasks was devised. Part training on those subtasks prior to the training of the task as a whole was compared to training of the task itself. Another comparison was made against an adaptive training regime, in which the time-pressure which is embedded in the task was reduced and then gradually increased according to the trainee's progress. Prior training on the subtask produced better performance. The results of the Adaptive Training indicate that there is a potential advantage in the use of this method but it has to be weighed against specific disadventages which arise from negative transfer.

The project has grown into a major research effort which will be conducted simultaneously in a number of Universities and research institutes. The topic of investigation common to all is training strategies which can be used in the teaching of a complex skill.

Prediction of Performance Quality in a Complex Perceptual Motor Task

Training pilots of tactical jet aircraft, or operators of complex dynamic systems, is time consuming and expensive. Consequently, the desirability of effective screening and selection instruments is well recognized. We are presently assessing the feasibility of using information from ERPs to aid in the prediction of performance on a complex, highly skilled task. Since ERPs provide information on cognitive processes not obtainable via behavioral measures, the likelihood of obtaining useful predictors is high.

We designed an ERP battery to provide information on a variety of ERP components in several experimental paradigms. We are currently examing the relationship between these data and performance in our criterion task, a complex and demanding arcade-like game, which we call "space fortress". In this game the subject must maneuver a space ship, identify and evade mines, and fire lasers to destroy the mines and, ultimately, a space fortress.

Forty right handed male subjects (ages 18-25) have completed four ERP sessions, followed by three sessions playing space fortress. In addition, a variety of psychometric instruments have been administered to each subject (e.g., Raven progressive matrices, symbol digit modalities test, a personality inventory). The ERP sessions include "oddball" paradigms, an S1-S2 go - no go choice RT paradigm, a Sternberg paradigm, and a dual-task tracking paradigm. These constitute our ERP battery. There is substantial research indicating that in these paradigms various ERP components (e.g., P300, the CNV) reflect the time necessary for stimulus evaluation and categorization, "context updating", resource allocation, and preparation for perceptual and motor processing.

Since we are also collecting traditional behavioral measures (e.g., RT and tracking performance), as well as information from the psychometric tests, we are able to evaluate the predictive utility of our ERP battery apart from the contribution of the more traditional measures.

Our criterion task, Space Fortress, was developed to reflect many of the cognitive components required for flying advanced jet aircraft. Given the value associated with an ability to accurately predict pilot performance, any increases in our predictive ability will be welcome.

References

Gopher, D. A selective attention test as a predictor of success in flight training. Human Factors, 1982, 24, 173-184.

Hillyard, S.A. and Hansen, J.C. Attention: Electrophysiological approaches. In M.G.H. Coles, E. Donchin and S.W. Porges (Eds.), Psychophysiology: Systems, Processes and Applications. New York, Guilford Press, 1984.

Isreal, J.B., Wickens, C.D., Chesney, G.L. and Donchin, E. The event-related brain potential as an index of display monitoring workload. Human Factors, 1980, 22, 211-224.

Kahneman, D. Attention and Effort. Englewood Cliffs, New Jersey, 1973.

Kramer, A.F., Wickens, C.D. and Donchin. An analysis of the processing demands of a complex perceptual-motor task. Human Factors, 1983, 25, in press.

Sternberg, S. On the discovery of processing stages: Some extensions of Donder's methods. Acta Psychologia, 1969, 30, 276-315.

Wickens, C.D. The structure of attentional resources. In R. Nickerson and R. Pew (Eds.), Attention and Performance VIII, Englewood Cliffs, New Jersey, Erlbaum, 1980.

Wickens, C., Kramer, A., Vanasse, L. and Donchin, E. Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information processing resources. Science, 1983, 221, 1080-1082.

PANEL SESSION

PANEL SESSION: AIR FORCE BASIC RESEARCH ON ABILITY MEASUREMENT

SESSION CHAIR: David L. Payne (Air Force Human Resources Laboratory)

PANELISTS: David L. Payne, William C. Tirre, Raymond E. Christal, Patrick Kyllonen (Air Force Human Resources Laboratory, Brooks AFB)

Gary L. Allen and Ben B. Morgan, Jr. (Old Dominion University)

PROCEEDINGS ENTRIES

"Individual differences in learning rate" (David L. Payne and William C. Tirre)

"Two for the money: Speed and level scores from a computerized vocabulary test" (Raymond E. Christal, William C. Tirre, and Patrick Kyllonen)

"Some people think faster: Can we measure it and who cares?" (Patrick C. Kyllonen)

"Assessment of learning abilities using rate measures" (Gary L. Allen and Ben B. Morgan, Jr.)

Individual Differences in Learning Rate

David L. Payne and William C.Tirre
Air Force Human Resources Laboratory

Abstract

This study examines the use of direct measures of learning rate as predictors of learning outcomes. Air Force recruits learned a list of name-number pairs to either 1, 2, or 3 successive error-free repetitions. After a break of several minutes, retention and ease of relearning to two perfect repetitions were measured. Results showed that initial learning rate was better than conventional intelligence scores in predicting retention and ease of relearning. Imposing a stringent learning criterion tended to diminish the differences between fast and slow learners in relearning only. The practicality of overlearning for slow learners is discussed.

As part of the Learning Abilities Measurement Program (Project LAMP) we are investigating the use of direct measures of learning rate as predictors of future learning and task performance. It is proposed that direct measures of learning rate collected under controlled laboratory conditions will help us to identify recruits who may score low on conventional tests, but have the potential to catch up with or surpass current high scorers in learning some new occupational skill. Given the present Air Force policy of assigning individuals to an occupational area on the basis of that area's rated learning difficulty, direct measures of learning rate could have great utility for future selection and classification procedures.

There are several unresolved issues concerning the use of measures of learning rate as predictors. In this paper we focus upon the use of learning rate as a predictor of the following learning outcomes: school achievement, knowledge re·ention, and ease of relearning.

The fir: issue is whether or not learning rate predicts future learning achievement :ter than conventional selection tests. Surprisingly few studies have addresse this issue. Gettinger and White (1979) measured learning rate as the number of trials required to reach a criterion level of performance on several academic learning tasks. Trials-to-criterion was a stronger correlate of school achievement than was a group administered intelligence test. Using a paired-associates task (nonsense syllables paired with either words or figures), Stevenson, Hale, Klein, & Miller (1968) measured learning rate as the amount learned in a controlled time period. Learning scores, as compared to IQ scores, correlated slightly better with school grades for male seventh graders, but the reverse was true for females. These studies suggest the relationship between learning rate and achievement will depend upon the kinds of learning task used, the achievement area predicted, and the population sampled in the study. In the near future, Project LAMP will explore this issue with both in-house and contractual efforts.

Another important learning outcome is knowledge retention. Do fast learners retain more (forget less)? This issue has received more attention, though no final verdict has been reached. Underwood (1954) argued that if fast and slow learners are brought to the same strength of association, there

should be no difference in retention. Early studies (Gregory and Bunch, 1959; Stroud and Schoer, 1959) reported slight advantages in retention for the fast learner. Later studies which equated level of learning in fast and slow learners via speeded versus unspeeded study times (Shuell and Keppel, 1970; Gentile et al., 1982) reported no differences in retention over repeated tests for the two groups. However, with academic learning tasks Gettinger (1983) found a correlation of -.81 between time-to-reach 100 percent mastery and delayed recall. Again, we lack closure on an important issue.

Highly related to knowledge retention is ease-of-relearning. The issue is whether or not students who learn fast initially also relearn fast after forgetting occurs. Individual variation in ease of relearning has not been researched thoroughly. Most attention has been on the effects of overlearning. For example, Schendel and Hagman (1980) found that a group of Army reservists who had been trained to a mastery level equal to twice the number of trials needed to attain proficiency, subsequently needed fewer trials and committed fewer errors in relearning the task back to proficiency. These results corroborate the traditional finding that degree of learning directly influences retention (Slamecka & McElree, 1983).

The interesting questions that emerge are (a) whether fast learners retain more and relearn more quickly; (b) whether individual differences in retention and relearning can be minimized by overlearning; and (c) whether general intelligence scores can substitute for direct measures of learning rate in the prediction of these learning outcomes. We addressed these questions in an experiment in which recruits first learned a list of name-number pairs to one of three degrees of learning, then took a delayed retention test and relearned the list.

Method

Subjects

The subjects in this study were 226 recruits in their sixth day of basic training at Lackland AFB, Texas. All subjects were tested in groups of 30 at individual testing stations. Each test station was equipped with a TERAK 8510A microcomputer outfitted with two dual density disk drives for data storage and a keyboard for response entry. After being briefed in a classroom on the purposes of the testing, subjects were assigned to testing stations and were familiarized with the computer keyboard.

Experimental Design and Test Procedure

A paired associates learning task consisting of 13 name-number pairs was administered via computer. After a study phase in which each pair (e.g., JONES = 13) was presented for three seconds, subjects began the test phase. In this test phase, each stimulus term would appear without its response term (e.g., JONES = ?) and the subject had to type in the correct answer. Upon answering, the computer would respond with "The correct answer is _____."

The experimental variable was the degree of learning imposed. Subjects were randomly assigned to one of three different groups and required to achieve criterion levels of 3, 2, or 1 error-free, successive repetition(s) of each item. Pairs were dropped from the list as the criterion was reached. Upon completion of the learning task, subjects were told that they would move on to a new task after a short break. Time on break was either 6, 12, or 18

minutes and this was crossed with degree of learning. Subjects were instructed to remain quietly at their stations during the break. After the break, subjects began relearning the same list as before, this time to a criterion of two error-free successive repetitions of each item.

The dependent variables were retention, measured as percent recalled on the first pass through the relearning of the 13-item list; and ease of relearning, measured as the number of trials needed to reach criterion on the relearning task. The independent variable, initial learning rate, was defined as the percentage of pairs learned in the first pass through the 13-item list immediately after the study phase. Armed Services Vocational Aptitude Battery (ASVAB) General composite scores were obtained as a measure of general intelligence.

Results

As was expected, percent correct recall increased as degree of learning increased (M = 33, 59, 69 and SD = 21, 24, 21, for the 1, 2, and 3 correct repetition(s) conditions respectively). In a likewise manner, trials to relearn decreased (M = 116, 62, 46 and SD = 79, 40, 24, respectively). Recall for the three-correct-repetitions group was over twice that of the one-correct repetition group, and relearning trials less than half. Time spent on break had virtually no effect; the same amount of forgetting was found for each level. For this reason, means were collapsed across break times. Initial learning rate correlated .46 with later recall and -.37 with number of trials to relearn the list. General intelligence (ASVAB General) correlated .25 with recall and -.28 with the relearning score. Thus, it appears that initial learning rate is a better predictor of learning outcomes than is general intelligence.

The questions of interest for the regression analysis are: (a) Does recall increase and trials to relearn decrease as level of learning increases?, (b) Does increasing the level of learning required diminish differences in retention or relearning between fast and slow learners?, and (c) Does initial learning rate account for learning outcome variance unaccounted for by general intelligence?

As the regression summary in Table 1 shows, both recall and ease-of-relearning were enhanced by increasing the required level of learning. The answer to the second question is no for retention and yes for relearning. The interaction effect for degree of learning by learning rate (DOL X LR) is small but significant, accounting for only 2.9 percent of the relearning variance. The regression slopes relating relearning to initial learning rate are -198, -107, and -39 for the 1, 2, and 3 perfect repetition(s) groups respectively. This decrease in slopes indicates that the difference between fast and slow learners in trials needed to relearn becomes smaller if both groups are brought to a high level of learning.

TABLE 1
Summary of Regression Analyses for Recall and Relearning

SOURCE	RECALL R^2	df	F	RELEARNING R^2	df	F
Degree of Learning (DOL)	.017	2	4.0*	.055	2	10.5**
Time	.000	1	ns	.003	1	ns
ASVAB-G	.001	1	ns	.028	1	10.6**
Learning Rate (LR)	.084	1	40.0**	.088	1	33.6**
DOL x ASVAB-G	.010	2	ns	.005	2	ns
DOL x LR	.008	2	ns	.029	2	5.5*
Full Model:	R^2 = .546, F(9, 216) = 28.9**			R^2 = .436, F(9,216) = 18.5**		

N=226
* $p < .05$
**$p < .01$

The regression analysis also shows that learning rate is a better predictor of retention and relearning (β = .498 and -.514, respectively) than is intelligence (β = .086 and -.329). This would suggest that if our goal is to predict learning criteria, we need direct measures of learning ability in our selection devices.

Discussion

The results of this study suggest that fast learners retain more and relearn more quickly than slow learners. These findings appear to contradict Underwood and others who suggest that no differences in forgetting should occur if the groups were equated in original learning. A critic might argue that simply bringing all learners to 100 percent mastery does not guarantee equivalent levels of association for fast and slow learners. The same critic might be pacified by our finding that individual differences in relearning were reduced somewhat by increasing the original level of learning, i.e., forcing subjects to overlearn the list. However, even with this reduction, slow learners (bottom 10 percent) in the 100 percent overlearning group required 25 times the number of trials to relearn as did fast learners (top 10 percent)--a finding which suggests that overlearning may be a very expensive compensation strategy.

One possible effect of our method of equating original learning was that slow learners experienced more within-list interference. Slow learners had more trials intervening between early learned items and later recall. By the time the more difficult items were learned, the easy items may have been forgotten. For this particular experimental design, the faster one learns the less likely one is to experience interference.

The finding that learning outcomes are predicted better by learning rate than by general intelligence must be interpreted with some caution. Yet to be determined is how well learning rate generalizes across tasks. The predictive advantage could be lost. Rather than one general learning ability, we suspect that there are several, each involving related but distinctive component processes. As part of Project LAMP, we intend to identify the primary components of learning and to use this knowledge in devising new selection tests.

REFERENCES

Gentile, J.R., Monaco, N., Iheozor-Egiofor, I.E., Ndu, A.N., & Ogbonaya, P.K. (1982). Retention by "fast" and "slow" learners. Intelligence, 6, 125-138.

Gettinger, M. (1983). Effects of learner ability and instructional modifications on time needed for learning and retention. Journal of Educational Research, 76, 362-269.

Gettinger, M., & White, M.A. (1979). Which is the stronger correlate of school learning? Time to learn or measured intelligence. Journal of Educational Psychology, 71, 405-412.

Gregory, S.C., Jr., & Bunch, M.E. (1959). The relative retentive abilities of fast and slow learners. Journal of General Psychology, 60, 173-181.

Schendel, J.D., & Hagman, J.D. (1980). On sustaining procedural skills over prolonged retention intervals. (Tech Rep No. 129-8) Alexandria VA: US Army Research Institute for the Behavioral and Social Sciences.

Shuell, T.J., & Keppel, G. (1970). Learning ability and retention. Journal of Educational Psychology, 61, 59-65.

Slamecka, N.J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 384-397.

Stevenson, H.N., Hall, G.A., Klein, R.E., & Miller, L.L. (1968). Interrelations and correlates in children's learning and problem solving. Monographs of the Society for Research in Child Development, 33, (7, Serial No 123).

Stroud, J.B., & Schoer, L.A. (1959). Individual differences in memory. Journal of Educational Psychology, 50, 285-292.

Underwood, B.J. (1954). Speed of learning and amount retained: A consideration of methodology. Psychological Bulletin, 51, 276-282.

AD-P003 339

TWO FOR THE MONEY--SPEED AND LEVEL SCORES FROM A COMPUTERIZED VOCABULARY TEST

Raymond E. Christal
William C. Tirre
Patrick C. Kyllonen
Air Force Human Resources Laboratory

Abstract

A speeded two-alternative computerized vocabulary test was constructed which yields both speed and level scores, but which takes considerably less time for administration than the ASVAB Word Knowledge test (7 vs 11 minutes). The level scores from the vocabulary test yielded higher validities than the Word Knowledge test for five paired-associates learning tasks included in an experimental battery, and the vocabulary speed (latency) scores proved to be useful in predicting individual differences in performance deterioration under conditions of information overload.

Introduction

As we enter the information age, many jobs are gaining new requirements for information generation, storage, retrieval, processing, and transmission. Even the pilot is becoming less an operator and more an information processor. There are two important aspects of information processing: accuracy and speed. We can do a pretty good job of predicting accuracy under normal circumstances with our present conventional tests. However, there has been no systematic effort by test psychologists to measure individual differences in information processing speed. Yet the ability of an individual to maintain processing accuracy under conditions of high information flow is critical in many military jobs during periods of emergency.

The concept of measuring information processing speed is not new. It has been a standard procedure employed by cognitive psychologists in laboratory experiments for the past two decades. However, most of these studies have dealt with group data and have been oriented toward the development and evaluation of theories relating to general, invariant mental processes and structures. Test psychologists also have attempted to extract both level and speed scores from conventional paper-and-pencil tests by using total test time and the number of correct responses. However, as Lohman points out (1979, pp 186-7), these studies are flawed. "Either the level scores are invalid because items are too simple, . . . or the speed scores reflect time for guesses, abandonments, and incorrect responses as well as time for correct responses." The availability of microcomputers makes it possible to obtain response latencies for only those items answered correctly, which suggests that we should now be able to derive meaningful information processing speed scores from conventional tests. The present paper discusses a first attempt to accomplish this. A vocabulary test was selected for the experiment since the items are homogeneous with respect to processing steps, and since a large variety of processing tasks contain a verbal component. A

set of elementary cognitive tasks was also included in the experimental battery to relate the verbal speed scores to scores from standard information processing tasks. Finally, a series of learning tasks with varying associated learning times was included to determine whether the verbal speed scores predict deterioration in performance as a function of learning speed requirements.

Method

Description of the Tests and Administration Procedures

Six predictor and five criterion tests were constructed for computer administration. The predictor set included a 94-item vocabulary test and five elementary information processing tasks: (a) simple reaction time; (b) two-choice reaction time; (c) physical identity; (d) letter identity; and (e) category identity. Stimuli for all of the predictor tests were presented on the screen in the form of a triangle, with a stem word on top and two response alternatives on the right and left below. Subjects were requested to respond as fast as they could without making mistakes, and to use the "D" key to select the left alternative or the "L" key to select the right alternative. The only exception was for the simple reaction time test, in which the "L" key was used to respond to the first detection of light on the screen. The criterion tests were five word-word paired-associates learning tasks, differing only in the exposure times for the learning stimuli, which were 8-, 4-, 2-, 1-, and .5-seconds, respectively. Each of these 40-item tests was made up of four 10-item learning and recall blocks. Half of the 5x4=20 blocks were administered at the beginning of the testing session, and the other half at the end. The blocks and items within blocks were assigned randomly to subjects. The tests between the beginning and ending learning blocks were administered in random order. /The entire battery was administered to 589 airmen in their sixth day of basic training at Lackland AFB, Texas, during the months of November and December 1983. Subtest scores from the Armed Forces Vocational Aptitude Battery (ASVAB) were retrieved from operational files and included in the data analyses.

Description of the Analyses

Microcomputer administration of the vocabulary test made it possible to compute the mean response latency for each subject using only those items answered correctly. However, there are two problems with using such scores as measures of verbal processing speed. First, there is a 50-50 chance of getting an item correct by guessing. Second, individuals with low scores correctly processed only the easy items, while individuals with high scores correctly processed both the easy and the difficult items. In order to avoid these problems, 40 very easy items were included in the test for computation of mean latency scores. Mean response latencies were also computed as speed scores for all of the information processing tasks. Percent correct scores were computed for all tests, including the learning tasks, as indicators of processing accuracy. Corrections for guessing were not necessary, since all items in all tests were completed by all subjects. Means, standard deviations, and intercorrelations were computed for all scores. Reliability estimates were computed using the odd-even split-half correlation technique, corrected to double length by the Spearman-Brown formula.

Results

The primary analysis results are presented in Table 1. It can be seen that, as predicted, performance on the learning tasks deteriorated as a function of stimulus exposure times; the mean accuracy being nearly 79 percent at the 8-second rate, and only 37 percent at the .5 second rate. Since the learning tasks contained 5-alternative items, the mean accuracy score for the .5-second level would have been only 19 percent if corrections had been made for guessing. This indicates a moderate floor effect for this task. The reliability coefficients show a modest drop-off at the 1-second rate, and a substantial drop-off at the .5 second rate, suggesting the need to correct for attenuation in order to avoid distortions in the predictive validity patterns.

Table 1: Validities of Selected Predictors for Five Learning Tasks

Learning Task Criteria				Predictor Validities			
				1	2	3	4
Exposure Time	M	SD	r_{11}	Vocab Latency	Vocab Level	Multiple 1&2	Word Knowledge
0.5	37.5	12.4	.62	-.42	.42	.52	.41
1.0	51.9	17.3	.81	-.33	.41	.47	.36
2.0	66.6	18.5	.86	-.24	.37	.40	.34
4.0	73.9	18.8	.90	-.19	.29	.31	.27
8.0	79.0	17.3	.89	-.14	.30	.31	.27

NOTES: Predictor validities corrected for attenuation in the criteria.

Table 1 reveals that the vocabulary latency scores added validity to the vocabulary level scores in predicting the learning task criteria, especially at the 1- and .5-second stimulus presentation rates. The verbal scores from the experimental vocabulary test had higher validities than the ASVAB Word Knowledge test for all of the learning tasks. The combination of vocabulary level and speed scores was a substantially better predictor of learning performance at the 1- and .5-second exposure rates than the Word Knowledge test (.52 vs .41 and .47 vs .36). This is especially encouraging in light of the fact that the experimental vocabulary test took considerably less time to administer than the ASVAB Word Knowledge test. The mean latency for items in the experimental vocabulary test was 1.931 seconds with a standard deviation of .9045 seconds. Allowing two standard deviations above the mean for slow subjects, and adding .5 seconds for inter-trial times, it is estimated that the 94-item vocabulary test for the type described in this study can be administered in about 7 minutes. This is a little more than half of the time required for the ASVAB Word Knowledge test, and only slightly more than that required for administration of a Word Knowledge test in the computer adaptive testing (CAT) format.

Table 2 reports the correlations among the vocabulary level, vocabulary latency and standard information processing task scores. Vocabulary level and vocabulary latency correlated only -.27, indicating that individuals with high vocabulary scores do not necessarily process vocabulary items quickly. With the exception of the category and letter identity tasks, none of the information processing tasks had a significant relationship with vocabulary level. On the other hand, all of the information processing tests correlated

significantly with vocabulary latency. The category identity test had the highest relationship (.754), which was as high as could be obtained from a multiple correlation of all five tests (R2.34567 = .754). It appears, then, that the processes measured by the vocabulary latency scores are very similar to those measured by the category identity task.

Table 2: Reliability Coefficients and Intercorrelations of Vocabulary Latency, Vocabulary Level and Five Information Processing Test Scores

Variable #	NAME	1	2	3	4	5	6	7	r_{11}
1	Vocab Level	----	-.27	-.16	.06	-.04	.04	.03	.87
2	Vocab Latency	-.27	----	.75	.56	.36	.34	.19	.97
3	Category Identity	-.16	.75	----	.72	.48	.45	.27	.98
4	Letter Identity	-.06	.56	.72	----	.58	.52	.29	.98
5	Physical Identity	-.04	.36	.48	.58	----	.39	.21	.99
6	2-Choice Reaction	.04	.34	.45	.52	.39	----	.38	.98
7	Simple Reaction	.03	.19	.27	.29	.21	.38	----	.98

The ASVAB Word Knowledge test (not shown) correlated .75 with the experimental vocabulary level, which is reasonably high, but lower than would be expected from virtually alternate forms. Inclusion of the vocabulary latency scores in a two-variable raised the correlation to .81. Vocabulary latency correlated -.48 with the Word Knowledge test, which is somewhat surprising considering the fact that vocabulary latency only correlated -.27 with vocabulary level. This may be an indication of a speed-accuracy trade-off factor operating in the vocabulary test.

Discussion

This study has demonstrated how latency scores obtained from a computer-administered vocabulary test can be useful in predicting individual differences in performance deterioration which takes place as a function of increases in information processing speed requirements. This finding has implications for selecting personnel for positions which demand the maintenance of processing accuracy in high information flow environments. The experimental vocabulary test took considerably less time to administer than the ASVAB Word Knowledge test (7 vs. 11 minutes), and yielded level scores having higher validities for all five learning tasks included in the experimental battery. As long as a subset of 40 easy items are included for computation of the latency scores, there is no reason why the rest of the vocabulary test could not be administered in the CAT format, thus having the best of both worlds. It also might be possible to compute latency scores from complex items if some method could be developed for eliminating items answered correctly by guessing. In a two-alternative item test, the expected value for the number of items answered correctly by chance is equal to the number of items answered incorrectly. The authors are presently investigating schemes for identifying the particular items which must be eliminated for each subject in order to compute latency scores for correctly processed items. If this is successful, the second step will be to use individual regression equations to rescale mean latency scores for all subjects to a common item difficulty level.

Results from this study suggest that latency scores may be uniquely important for predicting performance in situations where speed is required. Additional studies using a variety of materials and performance criteria are planned in order to fully investigate the utility of latency scores as predictors of information processing efficiency.

References

Lohman, D.F. (1979). Spatial ability: A review and reanalysis of the correlational literature. (Technical Report No. 8). Standford, CA: Stanford University, Aptitude Research Project, School of Education, October.

SOME PEOPLE THINK FASTER: CAN WE MEASURE IT AND WHO CARES?

Patrick C. Kyllonen
Air Force Human Resources Laboratory

Abstract

Many Air Force occupations require high speed filtering of incoming information and fast actions based on the information. Air traffic controllers are constantly engaged in situation evaluation and fast decision-making. Key-punch operators and other clerical workers scan numbers and letters and punch them extremely rapidly. Administrators scan reports, looking for relevant data, then make quick decisions based on what they see. In all these examples, there are tremendous differences between people in how quickly and accurately they are able to accomplish their tasks, even when the level of experience and training is held constant. For optimal manpower utilization, the Air Force must assign the fast scanners and decision-makers to jobs demanding those skills. In the current selection and classification system an individual's aptitude for these kinds of occupations is presumed to be indicated by his or her score on the Administrative composite in the ASVAB. The problem with this score, as we have determined empirically, is that it largely reflects speed of filling in answer sheets and only to a small extent reflects other components of high speed thinking that enter into most occupations. In this study we present a method for measuring directly the various components of high speed thinking and acting, and suggest some possible relationship to occupations that require them.

Assessment of Learning Abilities using Rate Measures

Gary L. Allen and Ben B. Morgan, Jr.
Old Dominion University

Abstract

Provided in this presentation is a description of progress in a unique research effort examining the utility of learning rate measures as predictors of learning abilities. Complex Experimental Learning Tasks (CELTs) have been developed in response to the challenge of devising a technology for measuring learning rate. CELTs provide a real-time sample of learning performance on criterion-free, face-valid memory tasks. Microcomputer-based administration affords detailed records of learing activities while facilitating efficient data management. Subjects' performance is described by plotting performance indices (e.g., accuracy, speed of responding) over time. Slope and intercept parameters from these functions hold promise as valid measures of learning rate. Efforts are underway to (a) determine relationships between traditional aptitude measures and learning rate measures from CELTs and (b) explore the predictive value of these rate measures for classroom performance.

Introduction

The effort to develop a technology for using learning rates to predict learning abilities is one thrust in a multifaceted research program aimed at advancing the state-of-the-art in ability testing. The goal to which this research program is dedicated is that of providing the empirical foundation necessary to update and improve the instruments used for military personnel assessment, classification, and selection. Within this context, the present research represents work toward the objective of combining microcomputer technology with an information processing approach to cognitive ability testing to arrive at a new perspective on the chronic issue of how best to assess an individual's capacity to learn.

The measurement of learning abilities has obvious implications for effective personnel decision-making in military settings. To date, learning abilities have been inferred from performance on pencil-and-paper tests measuring knowledge and cognitive skills acquired prior to the time of testing. Christal (1979, 1981) has questioned the efficacy of this practice and has suggested that rate of skill or knowledge acquisition might be a superior predictor of learning abilities. Because learning during training is a dynamic rather than static process (in that it involves change over time), it is reasonable to posit that a dynamic measure of learning might have advantages over aptitude test scores as predictors of future training success.

The advantages of, as well as the problems inherent in, measuring rate of learning have been well-documented (see Cronbach & Snow, 1977). The development of Complex Experimental Learning Tasks (CELTs) represents an effort to obtain these advantages and avoid these problems. CELTs are computer-administered tasks yielding a sample of learning performance. Each CELT is designed to (a) demand multifaceted processing in the sense of involving different cognitive factors or multiple cognitive operations; (b) place heavy demands on information encoding, storage, and retrieval; (c) impose no a priori performance criterion; (d) permit asymptotic performance to be achieved while allowing individual differences in final performance level; (e) bear some resemblance to generic real-world learning tasks while minimizing the influence of specific knowledge acquired prior to testing; (f) yield acquisition curves in less than one hour of administration time; and (g) readily lend itself to computerized administration to populations varying greatly in test-taking sophistication.

Underlying this approach to the measurement of learning rate is the assumption that learning abilities are actually combinations of cognitive skills and activities pressed into service in response to task demands rather than specific entitites per se. Accordingly, a number of different CELTs have been devised covering a range of processes from numerical operations through spatial orientation. In addition, it is assumed that a clearer understanding of these skills and activities is obtained when they are sampled in combinations rather than assessed with factor -referenced aptitude or ability tests.

CELTs have been used to measure learning rate in three experimental studies at the Center for Applied Psychological Studies at Old Dominion University. The first of these was to determine the extent to which rate of learning early in a CELT could predict asymptotic levels of performance in that task and in other CELTs, both in an original testing session and in a second testing session two-weeks later. The second study involved determining the relationship between performance on CELTs and performance on traditional pencil-and-paper tests of cognitive abilities. The third study was designed to explore the predictive utility of learning rate measures from CELTs for performance in training-oriented academic classes. Only the first of these three studies has been completed (Allen, Secunda, Salas, & Morgan, 1982); additional information regarding this study is provided in the following Method and Results sections.

Method

Subjects

Data were collected from 73 male subjects ranging in age

from 18 to 25 years, none of whom had more than one year of college education. Each subject received payment for approximately eight hours of participation.

Apparatus

CELTs were programmed and administered on Terak 8510 microcomputers, which feature standard LSI-11 microprocessors and dual eight-inch disk drives. Tasks were developed for administration on unmodified keyboards. Standard twelve-inch monitors provided 320 x 240 resolution. All programming was done in Pascal using UCSD-Pascal Operating System (Version II.0). Each microcomputer was located in a separate testing cubicle.

Learning Tasks

Four CELTs were used in the study. Each task was presented in an introduction phase, in which instructions were provided, and a test phase, in which items were presented continuously for 50 minutes. Practice items were presented in the introduction phase, and feedback with regard to accuracy and response time was provided after each item in the test phase. The CELTs included (a) the Coded Messages task, which required subjects to learn associations between words and symbols and to determine whether messages written in symbols were the same as verbal messages; (b) the Emergency Procedures task, in which subjects learned an ordered series of steps in an imaginary emergency situation and determined the accuracy of statements regarding the ordinal relationships among steps; (c) the Time Check task, which involved comparing digital time presentations with analogue time presentations while accounting for time zone differences; and (d) the Security Check task, which required learning a route through an imaginary military base and later answering questions about the location and security designation of various features.

Procedure

A repeated measures design was employed in which subjects were tested in two sessions separated by 10 to 14 days. In the initial session, subjects were administered the introduction and test phases of the Coded Messages, Emergency Procedure, and Time Check tasks. In the second session, they were administered the test phase of the Coded Messages and Emergency Procedure tasks and the introduction and test phases of the Security Check task.

Results

The primary analytic objective was to determine the extent to which subjects' learning rate early in the Coded Messages and Emergency Procedure during the first sessions could be used to predict asymptotic levels of performance in all four CELTs in

both sessions. To achieve this objective, several measures were derived from subjects' acquisition and re-acquisition curves, which were constructed by plotting the net correct responses (i.e., the number of items answered correctly minus the number answered incorrectly) per minute of testing time. Initial learning rate was represented by the slope and intercept of the initial 17 minutes (the first one-third) of this curve. The intercept reflected the subjects' ability to profit from instruction; the slope reflected a gain in accuracy and speed over time. Asymptotic performance level was measured in terms of the average net correct responses per minute over the final ten minutes of testing in each session.

Multiple regression analyses revealed the combination of slope and intercept to be a significant predictor of asymptotic performance level, in both the initial (multiple R's ranging from .51 to .76) and follow-up (.38 to .55) testing sessions. As expected, multiple R's were highest for within-task predictions. The evidence indicated that fast learners tended to achieve higher levels of performance than did slow learners.

An additional finding of interest was that asymptotic level of performance on a CELT in the initial testing was a strong predictor of both (a) performance decrement between the end of session one and the beginning of the second session and (b) the initial level of performance in the second session. In short, the results indicated that despite the fact that those subjects who acquired more knowledge/skill in initial learning showed a larger inter-session decrement than did those acquiring little knowledge/skill, higher achievers began performing in the second session at a level above that of lower achievers.

Discussion

The results of this exploratory study do not provide the empirical base necessary to revolutionize the field of aptitude testing. However, in demonstrating how computer-based technology and the process-oriented approach to cognitive assessment can be married in attempts to develop new approaches to old problems, these findings at least contribute to the revolution that is already underway. Clearly, the value of CELTs, as well as the concept of using rate of learning to predict learning abilities, will be determined only after more research is dedicated to standardization and validation.

In this regard, two studies are currently in various stages of completion. One of these efforts has involved determining the relationship between performance on CELTs and performance on a battery of pencil-and-paper tests of cognitive abilities. In this study, subjects completed a battery of CELTs (either the four described in this report or three additional tasks featuring

verbal, numeric, or spatial components) and tests of verbal, numeric, spatial, and memory abilities. The remaining study entails the prediction of classroom performance in training-oriented academic classes (e.g., engineering) from learning rate measures from CELTs. Results from these studies will help to determine the utility of this research thrust.

Acknowledgements

This research was carried out at the Center for Applied Psychological Studies at Old Dominion University under U. S. Air Force Contract No.F41689-83-C-0016. The research team included E. Pitts, R. Jones, E. Salas, M. Secunda, A. Gupta, J. Huskey, and A. Lea. The contract is monitored at the Test and Training Research Branch of the Air Force Human Resources Laboratory's Manpower and Personnel Division. The advice of Dr. Patrick Kyllonen, contract monitor, is gratefully acknowledged.

References

Allen, G. L., Secunda, M. D., Salas, E., & Morgan, B. B., Jr. (1982). Evaluation of rate parameters as predictors of the acquisition, decay, and reacquisition of complex cognitive skills (Tech. Rep. No. ITR-82-27). Norfolk: Old Dominion University, Center for Applied Psychological Studies.

Christal, R. E. (1979, October). What is the value of aptitude tests? Paper presented at the Annual Conference of the Military Testing Association, Gulf Shores, AL.

Christal, R. E. (1981, May). The need for laboratory research to improve the state-of-the-art in ability testing. Paper presented at the National Security Industrial Association-DoD Conference on Personnel and Training Factors in Systems Effectiveness, San Diego.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods. New York: Irvington Publishers.

PANEL DISCUSSION

MAN IN SPACE: SPACE TECHNOLOGY AND THE BEHAVIORAL ASPECTS

Ted Schroeder
Air Force Space Technology Center

John W. Ord and Ralph Lucianni
Headquarters AMD

Daniel H. Spoor
Headquarters Space Command

Len Gardner and Gary Coulter
NASA-Johnson Space Center

William L. Derrick
U.S. Air Force Academy

SESSION ABSTRACT

Military man-in-space is a reality. An Air Force individual has been scheduled as a Manned Space Flight Engineer to attend a military payload. Aerospace Medical Division and Space Division are exploring man's utility in military space systems. Space/Command is working on mission definitions and these, in turn, will determine man's application to the systems design to effect those missions. This panel has been designed to discuss these aspects in a dynamic, flexible manner. The panel will be sensitive to audience interaction as it seeks to develop an understanding of the mission and support of man in military space systems.

The U. S. military community has been exploring various manned military space vehicles such as the transatmospheric vehicle (TAV). These future vehicle concepts have military mission-unique crew requirements that neither the military nor NASA have addressed adequately to date. The TAV may require quick response deployment and long term presence of military crews in space for effective space-based C^3 systems. The C^3 architecture for conflict management involves man-machine integration and has, in fact, mandated studies by Electronic Systems Division (ESD) and Aerospace Medical Division (AMD). Military biotechnology research is required for optimizing vehicle design, crew station development, and manned performance in space.

The panel will emphasize the two major functional areas, man-machine integration and crew protection:

a. The objective on man-machine integration would address enhancement of man-in-the-systems, both when grounded and space based.

b. With respect to crew protection, the problems of additional stressors (e.g. accelerations, information displays, conflict management systems) might also be addressed.

Characteristics of Navy Recruits on Reading Comprehension and Related Variables: A Quarterly Three-Year Profile for FY 1981-83

Norman J. Kerr
Chief of Naval Education and Training
Harry L. Bowman
Memphis State University

Abstract

The purpose of the study was to describe the Navy's recruit population on reading comprehension and related variables by calendar quarter during the 3-year period of Fiscal Years 1981-83. Quarterly computer reports were the source of data on educational background, aptitude, and reading comprehension level of 233,749 recruits. The data revealed that approximately four fifths of the subjects are high school graduates while the typical recruit has completed slightly less than 12 years of school and reads near the mid-11th grade level. A significantly higher percentage of high school graduate recruits (3.73%) than non-high school graduate recruits (2.47%) are deficient in reading comprehension based on the 6.0 reading grade level (RGL) criterion. Recruits who score above the 6.0 RGL have a higher aptitude score mean than recruits who score below the 6.0 RGL. Variability on educational background and reading comprehension level was observed when the data on recruits were compared across the 12 quarters of the 3-year period.

Introduction

In June, 1978, the Secretary of the Navy issued a policy statement that addressed the need for training in basic educational skills among some personnel in order for the Navy to achieve its mission: "It is the policy of the Department of the Navy to provide, when indicated, on-duty remedial job-relevant basic skills training in order to fill personnel requirements" (SECNAVINST 1510.3, 1978, p. 1). The definition of job-relevant basic skills is "those capabilities prerequisite to efficient completion of Navy training programs, adequate job performance and career advancement. Notable among these are literacy, study and computational skills" (p. 1). In addition to new programs being developed as a result of the policy statement, fundamental skills training programs that were already operational were given official recognition. One of the operational programs, Academic Remedial Training (ART), provides instruction for recruits in basic reading and verbal skills needed to complete the demanding academic requirements of recruit training.

The ART Program was initiated officially by the Naval Technical Training Command in 1967 at the Navy's three Recruit Training Centers as a remedial reading course. The courses at the three sites evolved with highly dissimilar curricula and testing plans. The Chief of Naval Technical Training standardized the reading curriculum and testing procedures of ART and implemented a prescribed curriculum in 1978 (CNTECHTRAINST 1540.42). A verbal skills curriculum that focuses on listening and speaking skills primarily for recruits who speak English as a second language was added to the program at the Orlando and San Diego Recruit Training Centers in 1981.

When the standardization of the ART reading course occurred, a requirement was introduced that involved use of a Computer-Managed Instruction (CMI) system to perform test scoring functions and to maintain records containing data on recruits. The records that are created are used to prepare monthly and quarterly CMI reports on the educationally related and reading comprehension characteristics of the total recruit accessions for the period.

The problem defined for this study was to investigate the educationally related characteristics and relationships among selected variables that describe the Navy's recruit population by quarter during Fiscal Years 1981-83. The variable of primary interest was the recruit's grade level score on reading comprehension with secondary interest in years of education, aptitude score, high school graduation status, and quarter of entry into the Navy. The significance of the study is based on (1) identifying factors on which the recruit population varies across quarterly time periods and (2) demonstrating the use of CMI system data to monitor and manage the ART Program.

Procedures

The data utilized in this study were obtained from the quarterly CMI reports on Navy recruit characteristics and test performance. Data were available on 233,749 recruits who entered the Navy between October, 1980, and September, 1983. The reports provided data on four variables used in this study: years of education, high school graduation status, aptitude test score, and reading grade level (RGL) score on comprehension. A fifth variable, quarter of entry into the Navy, was defined by the quarter for which each report was prepared.

The Gates-MacGinitie Reading Comprehension Test (1978) and the Armed Services Vocational Aptitude Battery (1967-80) are the two tests for which data were reported. The Gates-MacGinitie, Levels D-F, is used as a screening instrument with all recruits who enter the Navy. (Only Level D was used prior to November, 1982.) The recruits who score below the 6.0 RGL on the first administration of the Gates are retested with an alternate form of Level D. (The data from the first administration of the Gates are used in this study.) Recruits who score below the 6.0 RGL on the second administration of the Gates are candidates for assignment to ART. The Armed Services Vocational Aptitude Battery (ASVAB) is administered to all recruits before they enlist in the Navy. The recruits included in this study had taken one of two series of ASVAB tests: Series 5-7 or Series 8-10. The aptitude score used in the study was the Armed Forces Qualification Test (AFQT) score. The AFQT is a normative percentile score that is derived from a composite of selected ASVAB scores.

The quarterly CMI reports provided means for years of education, AFQT scores of recruits classified by the 6.0 RGL criterion (above or below), and RGL scores. The median was calculated for RGL scores by quarter. Frequency distributions were compiled and percentage distributions were computed for selected data sets. Proportions were calculated for high school graduate and non-high school graduate recruits who scored below the 6.0 RGL criterion. The Chi-square procedure provided comparisons of frequency distributions for RGL scores, high school graduation status, and the 6.0 RGL criterion and high school graduation status by quarter of entry into the Navy. Confidence interval limits were computed to compare the proportions of high school graduates and non-high school graduates who scored below the 6.0 RGL. The .05 level of significance was used with the inferential analyses.

Results

The means on years of education were derived for recruits by quarter of entry into the Navy. The means varied from 11.6 years in the first, third, fifth, and seventh quarters to 12.1 years in the last two quarters. The mean for the 3-year composite group was 11.8 years of education. The data revealed a slightly increasing tread for quarterly means on years of education.

The AFQT (aptitude) percentile score means were obtained for recruits classified by 6.0 RGL criterion and quarter of entry into the Navy. Among the recruits with RGL scores below 6.0, the quarterly means for the AFQT ranged from 30.9 (Oct.-Dec., 1981) to 42.3 (Oct.-Dec., 1982). The AFQT quarterly means for recruits with RGL scores above 6.0 varied from 56.8 (Oct.-Dec., 1981) to 61.9 (Apr.-June, 1983). The means of the two composite groups were 35.8 for the recruits who had RGL scores below 6.0 and 59.1 for the recruits who had RGL scores above 6.0. While a pattern of change was not observable, the differences in the quarterly means for the high and low RGL groups were generally in the range of 20-25 percentile points. These differences suggest that the two RGL groups are relatively distinct with regard to AFQT scores.

The means and medians on RGL were determined for recruits by quarter of entry into the Navy. The means were relatively stable as evident by a range from 9.7 RGL (two quarters) to 10.0 RGL (four quarters) with a composite group mean of 9.9 RGL. The medians varied from 10.6 RGL (July-Sept., 1982) to 12.8 RGL (Apr.-June, 1983) while the composite group median was 11.4 RGL. After the introduction of multi-level Gates testing in November, 1982, the quarterly medians were consistently above the 12.0 RGL. The medians exceeded the means for all quarters, reflecting the negative skewness of the score distributions that is due to the population composition and the test norms.

The frequency and percentage distributions of recruits by high school graduation status and quarter of entry into the Navy were analyzed. The data for the composite group showed that 82.32% of the recruits were high school graduates. During the first seven quarters, the percentages were in the range of about 75-81%. In the next three quarters, the percentages were at the middle to upper 80% level. A temporary measure was instituted during June-September, 1983, that limited Navy enlistment, with few exceptions, to high school graduates. The effect was that 92.61% of the recruits in April-June, 1983, and 99.46% of the recruits in July-September, 1983, were high school graduates. The Chi-square comparison of the frequency distributions of recruits by high school graduation status and quarter of entry revealed that the distributions differed very significantly (Chi-square = 8,504.403, $\underline{p} < .001$).

High school graduate recruits were classified by the 6.0 RGL criterion and quarter of entry into the Navy to obtain frequency and percentage distributions. Within the composite group, 3.73% of the high school graduates scored below the 6.0 RGL on comprehension. The percentages varied by quarter from 2.54% (July-Sept., 1983) to 5.11% (Oct.-Dec., 1982). The quarterly percentages were essentially stable during the first eight quarters. After the highest percentage occurred in the quarter when multi-level Gates testing began, the percentages declined steadily during the last three quarters. The Chi-square analysis of the frequency distributions of high school graduate recruits classified by the 6.0 RGL criterion and quarter of entry indicated that statistically significant differences existed (Chi-square = 177.654, $\underline{p} < .001$).

The frequency and percentage distributions of non-high school graduate recruits were compiled by the 6.0 RGL criterion and quarter of entry into the Navy. The data for the composite group showed that 2.47% of the recruits who were not high school graduates obtained scores below the 6.0 RGL on

comprehension. Across the twelve quarters, the percentages ranged from 1.10% (Jan.-Mar., 1983) to 3.60% (Oct.-Dec., 1982). Cyclic increases in percentages occurred during the first four quarters and the next three quarters with a decrease between the fourth and fifth quarters. The percentages were erratic during the last five quarters. The results of the Chi-square comparison of frequency distributions for non-high school graduate recruits classified by the 6.0 RGL criterion and quarter of entry revealed significant differences (Chi-square = 48.951, $\underline{p} < .001$).

The recruits who scored below the 6.0 RGL criterion were classified by high school graduation status and quarter of entry into the Navy to derive frequency and percentage distributions. Based on the composite group data, 87.55% of the recruits with RGL scores below 6.0 were high school graduates. The quarterly percentages varied from 79.43% (Apr.-June, 1982) to 99.56% (July-Sept., 1983). While the quarterly percentages remained near or below 88% in eight of the first nine quarters, a consistent increase was evident during the last three quarters. The Chi-square comparison of the frequency distributions for recruits scoring below the 6.0 RGL criterion classified by high school graduation status and quarter of entry indicated that significant differences existed (Chi-square = 250.461, $\underline{p} < .001$).

Analyses were performed to compare the proportions of recruits who scored below the 6.0 RGL classified by high school graduation status and quarter of entry into the Navy. The data on the proportions of recruits scoring below the 6.0 RGL are discussed separately above as percentages for high school graduates and non-high school graduates. The relative proportions of the two groups who had RGL scores below 6.0 showed higher proportions for high school graduates than non-high school graduates in each of the twelve quarters and, consequently, during the 3-year period. The proportions of the two groups who scored below the 6.0 RGL were compared by computing confidence limits for the differences in proportions using the .95 level of confidence. The results indicated that the proportions for the two composite groups differed significantly because the confidence limits did not include 0 within the range. The proportions also differed significantly for 11 of the 12 quarters, excluding only the last quarter (July-Sept., 1983).

Conclusions

The Navy's recruit population is described in this study utilizing data on selected variables that apply to the recruits entering the Navy during a 3-year period of time. The conclusions are derived from the descriptive and inferential analyses of the data on these variables.

With regard to years of education, the typical Navy recruit has completed almost 12 years of education. Four fifths of the recruits are high school graduates. The inferential analysis on high school graduation status indicates that higher percentages of high school graduates are included among the recruits in the latter quarters of the 3-year period.

The aptitude of recruits as reflected by means on the AFQT does not appear to vary in any consistent pattern among the recruits with RGL scores below 6.0 or above 6.0. The means indicate that recruits who score above the 6.0 criterion have greater aptitude than recruits who score below the 6.0 RGL criterion. The difference in the aptitude score means for the two groups suggests that they are dissimilar populations.

The reading comprehension means of recruits exhibit relative stability across the 12-quarter period, placing the composite group near the 10th grade

level. The median reveals a disjuncture with the introduction of multi-level Gates testing in November, 1982. Prior to this time, the median was in the 10th to 11th grade range while it was in the 12th grade range afterward. The median for the 3-year period places the group near the mid-11th grade level.

Separate analyses of the distributions of high school graduate and non-high school graduate recruits classified by the 6.0 RGL criterion and quarter of entry indicate that differences exist across the 3-year period within each group. Among the high school graduates, the percentages who scored below the 6.0 RGL systematically decreased after the beginning of multi-level Gates testing while among the non-high school graduates, the lowest percentages occurred after initiating multi-level Gates testing.

Differences exist in the quarterly distributions of recruits who obtain RGL scores below 6.0 when classified by high school graduation status and quarter of entry into the Navy. Higher percentages of the recruits who scored below the 6.0 RGL were high school graduates in the latter quarters of the period. A higher proportion of the composite group of high school graduate recruits scored below the 6.0 RGL than was found in the composite group of non-high school graduate recruits. Significant differences exist for the composite groups and for the monthly groups in 11 of the 12 quarters.

This study indicates that the recruits who enter the Navy are likely to vary across fiscal year quarters on some important factors. As a result, ART management personnel need to plan for adjustments in the requirements for resources as the recruit population changes. By being aware of the variability in demands, more efficient allocation and utilization of resources can be achieved for ART. The study also demonstrates that information is available through the CMI reports on Navy recruit characteristics to improve the effectiveness of program management. Consistent monitoring of these reports and basic analyses of the data would enable the staff to determine the changes occurring in the composition of the recruit population that affect the ART Program.

References

CNTECHTRAINST 1540.42. (1978). Academic remedial training (ART). Millington, TN: Chief of Naval Technical Training.

MacGinitie, W. (1978). Gates-MacGinitie reading tests (2nd ed). Boston: Houghton-Mifflin Company.

SECNAVINST 1510.3. (1978). Remedial training in basic skills. Washington: Department of the Navy.

United States Military Enlistment Processing Command (1967-80). Armed services vocational aptitude battery (Series 5-10). Ft. Sheridan, IL: United States Military Enlistment Processing Command.

AD-P003 342

MMPI Normative Data For A Male Active Duty Army Population

Steven C. Parkison, Ph.D.
Walter Reed Army Medical Center

Francis J. Fishburne, Ph.D.
Walter Reed Army Medical Center

Abstract

The Minnesota Multiphasic Personality Inventory was administered to 1930 male active duty U.S. Army soldiers. After screening all subjects on a number of validity criteria, 1032 subjects were retained in the study and their data used to generate new normative tables for the MMPI. Since age and IQ levels have been found to significantly effect MMPI scores, the new normative tables are divided both by age and IQ. Preliminary work with the new normative tables is also presented.

Introduction

While the MMPI is likely the most widely utilized psychodiagnostic instrument with active duty military populations, little has been done to investigate the appropriateness of using the original MMPI normative data as a comparison group for military populations. Considering that the original normative data base for the MMPI was collected some 45 years ago (Dahlstrom, Welsh, and Dahlstrom, 1972) comparison of military MMPI scores with this data would appear, at least on the surface, to be inappropriate. Original normative data for the MMPI was collected in the late 1930's and early 1940's. While the data was refined some years later (Hathaway and Briggs, 1957) until very recently (Colligan, Osborne, Swenson, and Offord, 1983) there was virtually no contempory MMPI standards for normal individuals. Faschingbauer (1977) has suggested that while the original normative data was appropriate for use in the 1940's its usefulness in the 1980's is questionable.

Dahlstrom, Welsh, and Dahlstrom (1975), in reviewing research on the use of the MMPI, concluded that variations do appear to exist on the basic scales as a function of background variables such as age, sex, race and regional origin. The description of the normal 1940 Minnesota adult given by Dahlstrom et al. (1972) as being 35 years old, married, with eight years of schooling and living in a rural area stands in sharp contrast to the 25 year old, single, urban high school graduate who is the soldier of the 1980's. In addition, our society has dramatically changed in the last 40 years and these changes have had an effect on the individual's perception of his environment. Some behaviors that were considered pathological in the 1940's are today seen as appropriate and even healthy.

To date, MMPI research with military populations has focused almost exclusively on identified psychiatric populations. Only Bloom (1977) has examined a military population from a normative perspective. In his

study with U.S. Air Force basic trainees Bloom found a large number of subjects produced elevated MMPI profiles, but were reported to be free of psychopathology on subsequent interview by a psychologist or psychiatrist. While this data suggests MMPI scale differences between the Minnesota normative group and a military sample it has some limitations in applying the results to the military in general. First the Bloom (1977) data has a restricted educational and IQ range. Second, all subjects were in a stressful period, i.e., basic training, which could account for the elevated profiles. And finally, the Bloom data represents a rather restricted age sample since all subjects were in basic training.

Furthermore, MMPI researchers (Colligan et al., 1983) have begun to question the T score transformations utilized with the original data. A common assumption among clinicians is that the T scores on the MMPI represent a normal distribution of scores. This assumption is faulty in that the T score transformation in common use for the MMPI simply does not alter the underlying distribution, but rather preserves the skewed underlying distribution.

The purpose of this study was to obtain MMPI's from a representative sample of active duty military personnel and generate an appropriate normative base for the use of the MMPI with this population. Additionally, it was deemed paramount to develop transformations that would avoid the hazard of using simple linear transformations such as the T score transformation on non-normal distributions.

Method

Subjects

Subjects were 1930 male active duty U.S. Army soldiers stationed at various locations throughout the United States and Europe. All subjects were at least 18 years old, had one year or more of active duty service and had volunteered to participate in the study. Ninety percent of the subjects were enlisted and 10 percent were officers. Ages ranged from 18 to 32 with a mean age of 25 years and a mean education of 12 years. Job description was representative of the Army as a whole with 60 percent of the subjects being in support branches and 40 percent being in combat arms branches.

Procedures

Along with completing the MMPI subjects were administered the Shipley Institute of Living Scale Vocabulary and Abstraction Test. Subjects also completed a 43 item background information form. While the background information questionaire provided demographic data on the subjects, the Vocabulary and Abstraction Test yielded an IQ score (Boyle, 1967) Standard MMPI group form booklets were used and the answer sheets machine scored using an optical scanner. Only fully completed forms were utilized in the analysis and any partially completed forms were discarded along with the rest of the subjects' materials.

Results

The data were screened for validity using criteria suggested by Greene (1980) and Dahlstrom et al. (1972). An MMPI profile was considered valid and the subject eliminated from the sample if: (a) 30 or more questions were left blank, (b) scale F had 25 or more questions answered in a critical direction, (c) the Carelessness Scale (Greene, 1980) was 6 or higher. Furthermore, subjects with an IQ score below 75 were not included because of questions as to their ability to fully understand the questions and possible low reading levels. Several other criteria were also considered important in producing a normative population. Soldiers who reported a felony conviction or court marital were not included in the sample. Subjects endorsing a psychiatric hospitalization, suicide attempt, psychiatric treatment of any nature or treatment for drug or alcohol problem were also excluded from the sample. While no amount of screening can completely produce a "normal" population the above criteria were considered strenuous enough to produce a sample of relatively healthy, well functioning active duty soldiers.

A total of 898 subjects (46%) were eliminated from the study based on one or more of the screening criteria. This left 1032 subjects (54%) as the data base from which to develop new normative tables.

An examination of the distribution of raw MMPI scores revealed all scales to be significantly skewed, a finding consistent with that found on the original MMPI data gathered in the 1940's and also more recent data collected in Minnesota (Colligan et al., 1983). Since a simple linear transformation such as the T transformation would do nothing to normalize these distributions, a Box-Cox power transformation technique (Box and Cox, 1964) was utilized to find a transformation that would produce normal, Gaussian distributions of scores. Scale F was not transformed, but rather left skewed because the scale's original development was as an "infrequent item scale" thus a positive skew in the distribution of this scale was to be expected and was retained to avoid increasing false positive elevations. Table I summarizes the transformations used for each scale.

Development of Age and IQ Tables

Age and IQ levels have been found to have considerable effect upon the MMPI (Greene, 1980; Fishburne and Parkison, 1984). In view of these findings it was considered desirable to develop specific tables of norms by age and IQ levels. This would permit the clinician to not only compare an individual's MMPI scores to others in the Army, but also to compare the score to other scores from the appropriate age and IQ group, thus accuracy of description should be improved. Initially, subjects were compared across each age year. Examination of the data revealed that while there was very little difference across some age groups, others showed significant differences. Thus, age could be grouped based upon the points of significant variations. Ages were eventually grouped into 4 levels; 18-20, 21-23, 24-26, 27-33. The process for grouping IQ level was virtually identical and produced 3 distinct IQ groups; 75-98, 99-110, 111 and above.

Use of these normative tables is quite straight forward and involves knowing the subjects raw scores on the MMPI, age and IQ level from the Shipley Hartford Institute of Living Scale. The subject's age and IQ determine what table is appropriate and the MMPI raw scores of the subject are converted to the T-scores from the table.

TABLE I
Transformations and Summary Statistics

Scale (X)	Transformation (Y)	$\overline{Y}$	SD	Skewness Coefficient of Y	Percent with T 30	70
L	$Y=\sqrt{X+1}$	2.22	0.51	0.10	1.84	2.42
F	$Y=X$	9.07	5.83	0.75	4.36	0.00
K	$Y=X^{.8}$	7.43	2.33	0.11	1.45	1.36
1+K	$Y=\mathrm{Log}_{10}(X+1)$	1.15	0.14	-0.10	2.71	2.62
2	$Y=X^{.3}$	2.44	0.19	0.09	1.94	2.62
3	$Y=\sqrt{X+1}$	4.53	0.57	0.12	2.13	3.00
4+K	$Y=\sqrt{X}$	4.88	0.49	0.00	0.97	1.36
5	$Y=\sqrt{X}$	4.92	0.48	-0.02	4.36	2.42
6	$Y=(X+1)^{.4}$	2.72	0.36	-0.05	2.03	2.91
7+K	$Y=\sqrt{X}$	5.34	0.58	0.07	1.07	1.36
8+K	$Y=\mathrm{Log}_{10}(X+1)$	1.50	0.12	-0.11	1.55	1.45
9+K	$Y=X^{.8}$	12.67	2.04	-0.04	1.74	1.07
0	$Y=X^{.8}$	14.67	3.52	-0.07	1.84	2.81

Discussion

The development of these normative tables represents a considerable improvement in the MMPI data base available to clinicians working with active duty soldiers. There are several advantages to using the MMPI tables developed from this study. First, the clinician will be able to compare his patients' MMPI results with those of other active duty Army soldiers. Second, the clinician may now compare his patients' results with the results of other individuals of the same age and IQ level. Finally, since these tables are developed from normal distributions, inferences may more accurately reflect deviations from normal limits of psychological functioning.

Further work needs to be completed to develop female active duty norms and expand the current age level limitations in the data. Additionally both racial and regional differences on the MMPI scores appear to be fruitful areas of investigation.

Due to limitation in allotted space a copy of the normative tables was not included, but is available on request from the senior author.

REFERENCES

Bloom, W. Relevant MMPI norms for young Air Force trainees, Journal of Personality Assessment, 1977, 5, 505-510.

Butcher, J.N. & Tellegren, A. Common methodological problems in MMPI research, Journal of Consulting and Clinical Psychology, 1978, 46, 620-628.

Colligan, R.C., Osborne, D., Swenson, Wendell M. & Offord, Kenneth P. The MMPI; A Contemporary Normative Study. New York: Praeger Publishers, 1983.

Dahlstrom, W.G., Welsh, G.S. & Dahlstrom, L.E. An MMPI Handbook (Vol. 1). Minneapolis: University of Minnesota Press, 1972.

Dahlstrom, W.G., Welsh, G.S., & Dahlstrom, L.E. An MMPI Handbook (Vol. 2). Minneapolis: University of Minnesota Press, 1975.

Fashingbauer, T.R. The future of the MMPI. In C. Newmark (Ed.), MMPI clinical and research trends. New York: Praeger Publishers, 1979.

Fishburne, F. & Parkison, S., Age effects on active duty Army MMPI profiles. Paper presented at DOD Psychology Symposium, 1984.

Greene, R. The MMPI, An interpretive manual. New York: Grune and Stratton, 1980.

Hathaway, S.R., & Briggs, P.F. Some normative data on new MMPI scales. Journal of Clinical Psychology, 1957, 13, 364-368.

AD-P003 343

Age Effects On Active Duty Army MMPI Profiles

Francis J. Fishburne, Ph.D.
Walter Reed Army Medical Center

Steven C. Parkison, Ph.D.
Walter Reed Army Medical Center

Abstract

Age affects on response patterns to the Minnesota Multiphasic Personality Inventory have been recognized by the creation of separate norms for adolescent, adults and aged adults. This study examined the effects of age within a normal Army adult male population on MMPI response pattern. Of the validity and clinical scales only scales L, 3 (Hysteria) and 5 (Masculinity-Feminity) showed no significant age effects. Scale means and percentage of respondents scoring above 70T (non-K-corrected Minnesota Adult Norms) across age groups both showed scale age clusters which substantiate the need for age appropriate norms within the adult Army male population.

Introduction

MMPI literature suggests that age is an important influencing variable on MMPI response patterns. This study was designed to examine if age effects within the 18 to 33 year range are sufficiently variable to warrant construction of normative tables which take the age effect into consideration.

Gynther (1979) provides an excellent review of the literature examining age effects on the MMPI. Two studies, Swenson (1961) and Hathaway and Monachesi (1963) compared special age groups of normal subjects, aged and adolescent respectively, with the Minnesota adult normal group. Statistically significant findings of differences between several scale means of the special age groups as compared to the Minnesota adult normal group have led to the development of separate normative tables for adolescents and aged populations. Other studies using normal groups which were reviewed by Gynther examined differences between older and younger subjects and identified characteristics of older subjects. Gynther pointed out that although statistically significant differences were often found, the mean scale scores were not clinically significantly deviant from the Minnesota adult normal mean. Gynther did not define what he meant by clinical significance. Thus, Gynther appears to question if small statistically significant mean scale differences are meaningful in terms of clinical significance.

Within military populations evaluated in mental health clinics and hospitals it has been common to find elevated, 70 to 80T, MMPI profiles given by soldiers who have demonstrated reasonably adequate adjustment. Bloom (1977), using Air Force basic trainees, found a large number of airmen who produced highly elevated, up to 90T, MMPI profiles who, when given a mental status exam by a psychologist or psychiatrist, were judged

to have no pathologically deviant personality characteristics. Bloom administered the MMPI to a large sample of male and female airman trainees and found statistically significant differences between the mean scores of the trainees and the Minnesota adult normals. Using scale means provided in Bloom's (1977) study and the non-K-corrected Minnesota adult conversion table (Dahlstrom, Welsh, and Dahlstrom, 1975) only one scale for males, Ma, and no scale for females reached a T-score of 60 or higher.

Greene (1980) reviewed the studies related to age norms and reported that ninth graders (Hathaway and Monachesi, 1963) had mean K-corrected scores on scales 4, 8, and 9 which were approximately 10-T score points higher than Minnesota adult normals. Swenson's (1961) aged group had no mean K-corrected scores on any scale which deviated as much as 10-T score points above the Minnesota adult normals. Despite Gynther's implication that a statistically significant mean difference may not, in fact, result in clinically significantly deviant profiles, separate normative tables have been developed and published for adolescent and aged populations.

This study examines the relation between age and MMPI scale scores of normal Army soldiers who have made a successful adjustment to the military environment.

Method

Subjects

Subjects were 1930 male active duty U.S. Army soldiers stationed at various locations throughout the United States and Europe. All subjects were at least 18 years old, had one year or more of active duty service and had volunteered to participate in the study. Ninety percent of the subjects were enlisted and 10 percent were officers. Ages ranged from 18 to 33 with a mean age of 25 years and a mean education of 12 years. Job description was representative of the Army as a whole with 60 percent of the subjects being in support branches and 40 percent being in combat arms branches.

Procedures

Along with completing the MMPI subjects were administered the Shipley Institute of Living Scale Vocabulary and Abstraction Test. Subjects also completed a 43 item background information form. While the background information questionaire provided demographic data on the subjects, the Vocabulary and Abstraction Test yielded an IQ score (Boyle, 1967). Standard MMPI group form booklets were used and the answer sheets machine scored using an optical scanner. Only fully completed forms were utilized in the analysis and any partially completed forms were discarded along with the rest of the subject's materials.

Two-way analyses of variance were utilized to examine main and interaction effects of age and intelligence on the validity and clinical scales. In order to identify the number of respondents above 70-T (non K-corrected) using Minnesota adult norms, histograms showing the

distribution of scores on each scale by one year age intervals, 18 to 33, were generated. Scale means by one year age intervals, 18 to 33, were computed and converted to non K-corrected T scores for comparison across ages.

Results

The data were screened for validity using criteria suggested by Greene (1980) and Dahlstrom et al. (1972). An MMPI profile was considered invalid and the subject eliminated from the sample if: (a) 30 or more questions were left blank, (b) scale F had 25 or more questions answered in a critical direction, (c) the Carelessness Scale (Greene, 1980) was 6 or higher. Furthermore, subjects with an IQ score below 75 were not included because of questions as to their ability to fully understand the questions and possible low reading levels. Several other criteria were also considered important in producing a normative population. Soldiers who reported a felony conviction or court marital were included in the sample. Subjects endorsing a psychiatric hospitalization, suicide attempt, psychiatric treatment of any nature or treatment for drug or alcohol problem were also excluded from the sample. While no amount of screening can completely produce a "normal" sample the above criteria were considered strenuous enough to produce a sample of relatively healthy, well functioning active duty soldiers. A total of 898 subjects were eliminated from the study based on one or more of the screening criteria. This left 1032 subjects in the data base for this study.

TABLE 1

Two-way Analysis of Variance: Significant
F-Ratios for Age, IQ and Age x IQ Interaction

Scale	Age	IQ	Age x IQ
L	0.11	23.37***	1.09
F	25.51***	52.48***	0.43
K	4.67**	6.57**	1.66
1(Hs)	5.20***	33.86***	0.95
2(D)	2.94*	16.87***	1.71
3(Hy)	1.63	2.57	1.16
4(Pd)	12.69***	17.44	1.09
5(Mm)	0.83	17.84***	1.16
6(Pa)	11.91***	27.84***	0.42
7(Pt)	22.20***	31.11***	0.76
8(Sc)	35.10***	54.02***	0.42
9(Ma)	24.26***	20.05***	2.93**
0(Si)	4.13**	16.36***	1.03

* $p < .05$ ** $p < .01$ *** $p < .001$

Table I shows results of the two-way analysis of variance for age and intelligence. Interaction effects of age and intelligence were only significant on scale 9(Ma). Except on Scale 3, all main effects of

IQ were significant. This finding will be addressed in future studies. Main age effects which are significant were found on all scales except L, 3, and 5.

Table II

T Scores for Scale Means (Minnesota Adult Norms)
Non K-Corrected in the Army Sample

	18 (N=39)	21 (N=127)	24 (N=86)	27 (N=37)	30 (N=38)	33 (N=28)
L	50	50	53	50	50	50
F	70	66	64	60	60	60
K	48	49	49	53	53	53
1(Hs)	60	58	56	51	53	51
2(D)	60	58	56	58	58	58
3(Hy)	58	56	55	56	55	56
4(Pd)	68	65	63	58	58	58
5(Mm)	59	57	57	57	59	61
6(Pa)	65	62	62	56	56	56
7(Pt)	66	60	59	54	53	52
8(Sc)	78	67	63	57	55	55
9(Ma)	70	68	68	61	61	59
0(Si)	56	54	54	53	54	52

Mean scale score conversions to T-scores (non K-corrected) were computed for each scale by year group to determine the relative deviation of the soldier means from the Minnesota Adult norms. Table III shows the means for the year group at three year intervals between ages 18 and 33 (because of space limitations full data for tables II and III are not provided but can be obtained from the senior author). Scales 7, 1, 4, 6, 7, 8 and 9 all show a trend of higher to lower scale mean T scores from younger to older year groups. On several scales shifts in mean scale scores show relatively smooth linear changes across year groups, e.g. scale 7, and on others a somewhat sawtooth pattern with several year group means remaining the same or nearly the same before a drop in the mean occurs. e.g. scale 6.

Because the underlying distribution of scores in the Minnesota Adult normative table are not normally distributed, the distributions are skewed generally to the right (Colligan, et. al., 1983), is also important to review age effects on the number of respondents by year groups who fall beyond 70T on the Minnesota Norms. Table III shows the percentage of subjects by scale that fall beyond 70T on the scale in the Army sample for each third year group. As would be expected from Table II data, the percentages decline from younger to older year groups on scales F, 1, 2, 4, 6, 7, 8, and 9.

Discussion

The data in this study show a significant effect of age on MMPI response patterns within a normal group of Army males. The scale means and percentage of subjects who score above 70T when raw scores are corrected to T scores using non K-corrected Minnesota Adult norm tables

Table III

Percentage of Subjects in the Army
Sample with Non K-Corrected T Scores
(Minnesota Adult Norms) Over 70 T

	18 (N=39)	21 (N=127)	24 (N=86)	27 (N=37)	30 (N=38)	33 (N=28
L	0.0	0.0	1.2	2.7	2.6	0.0
F	38.5	29.1	20.9	16.2	13.2	14.3
K	2.6	0.0	3.5	0.0	2.6	0.0
1(Hs)	12.8	11.8	11.6	5.4	7.9	3.6
2(D)	25.6	18.9	11.6	13.5	13.2	17.9
3(Hy)	7.7	7.1	5.8	10.8	7.9	3.6
4(Pd)	33.3	33.9	22.1	5.4	15.8	17.9
5(Mm)	10.3	11.8	4.7	10.8	5.3	7.1
6(Pa)	20.5	23.6	16.3	13.5	5.3	0.0
7(Pt)	33.3	19.7	17.4	18.9	0.0	7.1
8(Sc)	61.5	36.2	24.4	18.9	10.5	17.9
9(Ma)	61.5	39.4	34.9	13.5	15.8	7.1
0(Si)	5.1	3.9	2.3	8.1	7.9	10.7

show patterns of differences across the year groups where younger year groups have clinically significantly more deviation than older year groups. A single combined normative table for this population would tend to overclassify in the younger year groups and underclassify in the older year groups.

The underlying distribution needs to be transformed to normal distributions and then clinically significant differences between year groups can be identified to arrive at the best year groups combinations for establishing Army normal tables (See Parkison and Fishburne, 1983).

REFERENCES

Bloom, W. Relevant MMPI norms for young Air Force trainees. Journal of Personality Assessment, 1977, 5, 505-510.

Colligan, R.C., Osborne, D., Swenson, W.M., & Offord, K.P. The MMPI; A Contemporary Normative Study. New York: Praeger Publishers, 1983.

Dahlstrom, W.G., Welsh, G.S., & Dahlstrom, L.E. An MMPI Handbook (Vol.2). Minneapolis: University of Minnesota Press, 1975.

Greene, R. The MMPI: An interpretive manual. New York: Grune and Stratton, 1980.

Gynther, M.D. Aging and personality. In J.N. Butcher (Ed), New developments in the use of the MMPI. Minneapolis: University of Minnesota Press, 1979.

Hathaway, S.R., & Monachesi, E.D. Adolescent personality and behavior: MMPI patterns of normal, deliquent, dropout, and other outcomes. Minneapolis: University of Minnesota Press, 1963.

Parkison, S. & Fishburne, F., MMPI normative data for a male active duty Army population. Paper presented at DOD Psychology Symposium, 1984.

Enlisted Performance Evaluation System for the U. S. Coast Guard

Vincent W. Patton III, Ed.D.
Petty Officer First Class, U. S. Coast Guard
U. S. Coast Guard Headquarters

Abstract

The U. S. Coast Guard has revamped its Enlisted Performance Evaluation System on 1 July 1983. The new system was developed by a five member project team which analyzed in detail the problems and shortcomings of the previous system, and re-defined needs and uses of the new system. The prime consideration in designing the new system was to better identify performance levels and provide feedback of strengths and weaknesses to the evaluee.

Introduction

On 1 September 1982, an effort began to revise the enlisted evaluation system which had served the Coast Guard for over 25 years. The system was in need of a major overhaul to keep up with the growing personnel management requirements and because evaluation scores were generally inflated. The shortcomings of the initial system were analyzed in detail and later helped define the goals and purposes of the new evaluation system. This report will provide a brief and comprehensive background on the process used to design the new Enlisted Performance Evaluation System, which affects over 40,000 active and reserve enlisted members of the Coast Guard.

Background

The initial evaluation system was developed in 1957. Prior to that the Coast Guard had no formal appraisal system for enlisted personnel. The primary purpose of the system was to "differentiate between the outstanding, the excellent, the average, the below average, and the unsatisfactory," which provided a numerical gauge of performance to support personnel policy decisions in areas such as advancement, special assignments, awarding the Coast Guard Good Conduct Medal, and eligibility for discharges and reenlistments.

The system measured work performance of all enlisted members from pay grades E-1 through E-9 on the same form, using three performance evaluation categories: Proficiency, Leadership, and Conduct. Proficiency measured the enlisted member's performance on the job, and was loosely defined as "...skill, efficiency, and knowledge of specialty, and demonstrated ability to perform effectively." Leadership (applied only to pay grades E-3 to E-9) measured the personal qualities which one should or does possess to perform managerial responsibilities, and was defined as "...ability to plan and assign work to others, and to effectively direct their activities, and his/her ability to maintain proper military relationships with other service personnel. Ability to recognize and carry out his/her civil rights/human relations responsibilities." Conduct measured the members ability to conform to rules, regulations, and military standards.

The system provided only one mark in each area, numerically indicated under a "four-point-O" (4.0) system. Proficiency and Leadership marks were frequently identical, and Conduct was almost always 4.0. Although the form provided "some" performance standard definition for each point on the scale, i.e.,"Grossly Inadequate" for 0, "Excellent" for 3.7/3.8, and "Ideal" for 4.0), evaluators were free to apply their own interpretations of these definitions. The form defined "average" performance as 3.3, but the servicewide view of "average" was in correlation to the pay grade (i.e. E-4 to 3.4, E-5 to 3.5, E-6 to 3.6, etc.) creating inflation.

As a result the system did not provide adequate differentiation between members of the same pay grades. Junior and senior enlisted members were evaluated on the same form at the same time. The system did not provide the evaluee with information on what his or her strengths/weaknesses were. Proficiency, Leadership, and Conduct were broadly defined, and all the evaluee was privy to was the "final" three marks, with some understanding that these marks would refer to terms such as "Outstanding, Excellent, or Average." The mark assigned in each category was an "average" of the items that made up that category. Therefore, the system did not provide any meaningful feedback on job performance, only final marks.

Discussion

Structuring the New System

After a good deal of study of the shortcomings of the system, and several discussions with the "users" of the system (personnel management policy decision makers), the structure of the new system began to take shape. By 1 December 1982, a outline plan for designing and implementing a revised system was submitted to the Chief, Office of Personnel with a planned implementation date of 1 July 1983. There were three specific purposes for the revised evaluation system:

(1) To capture a valid and reliable assessment of the performance of enlisted members to allow the Coast Guard to promote and assign them with a high degree of confidence, and to perform other necessary administrative actions such as reenlistments and discharges.

(2) To provide feedback to show each member how well he or she is performing in the areas measured. The evaluation will identify areas where each member can work to improve his/her own individual performance, and will provide details to support counseling whether requested by the individual or deemed necessary by higher authority.

(3) Emphasize important Coast Guard values.

The outline plan was broken down into three phases: design, implement, and monitor.

The views and ideas of the organization regarding enlisted performance appraisal as expressed in various study papers and memos indicated a desire to:

(1) Strengthen the ability to differentiate in order to identify high performers so as to better use and reward them, the middle performer so as to train and further develop them, and low performers for purposes of initiating corrective actions or discharge.

(2) Expand the role of the present system by identifying the needs of the individual and then develop a process that will, where possible, fulfill these needs.

(3) Have an enlisted evaluation system that is not intimidating or complex to use.

Coast Guard officers Stumpff and Chevalier prepared the first in-depth analysis of the enlisted evaluation system. In their 1976 graduate thesis they demonstrated statistically that the system was not adequately differentiating between the pay grades, and thus not supporting the intended uses of the evaluation system. They found fault with the evaluation process and noted that the system provided no feedback to the individual and had no monitor or control system. They advocated an approach that captured instances of behavior both good and bad by recording the information on a "Significant Incident Form," that could be used as the basis for a periodic evaluation and as the foundation of a performance feedback system. Stumpff and Chevalier's proposed system also called for the use of multiple evaluation forms, to provide differentiation of responsibilities between the pay grades.

In late 1976, to deal with the problem with a lack of differentiation between the enlisted pay grades, Ohio State University (Cornelius and Hakel, 1978) was awarded a contract to develop an improved enlisted personnel evaluation system. The study had four objectives:

(1) determine how many different evaluation forms there should be;

(2) develop prototypes of the proposed evaluation forms;

(3) develop system maintenance procedures;

(4) propose a field tryout of the new developed system.

The study concluded that a series of forms based upon job rating clusters (i.e. Aviation, Deck/Watch, Engineering, Electronics, etc) as determined by an extensive job element inventory were impractical and that a more satisfactory approach involved forms based upon pay grade groupings (non rated or E-1 through E-3, petty officer or E-4 through E-6, and chief petty officer or E-7 through E-9). They advocated an approach that began with a high quality measuring instrument, training the evaluator to improve his or her evaluation skills, and motivating the evaluator to do well by use of a monitor or control system. The recommended proposals from the results of the OSU study were not adopted by the Coast Guard.

Design and Implementation Phase

The design phase was broken down into identifiable tasks for each project member so as to build a scheduled and systematic work plan use the information from the OSU study, the graduate theses, and direction from the Commandant and the Chief Office of Personnel. An organizational requirement was that the design of the new system involve "Coast Guard participation." Competent officers and enlisted personnel were selected from field units. Their experience and input helped design the system and gave them some "ownership" which helped in the implementation process.

The selection of performance criteria and the development of performance standards by this group allowed the system to reflect the attitude of the working Coast Guard towards various aspects of performance appraisal. This was done by organizing three project field panels, one for each pay grade group: E-1 to E-3 (nonrated personnel), E-4 to E-6 (petty officers), and E-7 to E-9 (chief petty officers). Each panel developed personal performance characteristics and mixed anchored rating scales for the pay grade group they were assigned. This "critical incident technique" (Flanagan, 1954), resulted in a series of performance scales specific to the job for which they are created and retains all the terminology used by evaluators to describe the job. A similar method was used in developing the Coast Guard Officer Performance Evaluation System in 1980, and the Canadian Forces Performance Appraisal System in 1981 (Shields, 1982). The panel groups defined only three scales: Marginal, Middle, and High. Marginal describes performance that is "below the minimum acceptable level," middle describes "the Coast Guard's expected standards of performance," and high describes "the exceptional performer." The panel members were a cross section of subordinates and supervisors, as well as executive and commanding officers of various ranks and pay grades.

The three panels provided over 130 personal performance characteristics many of which were tautological, with the low, middle, and high rating scales defined for numerous categories. After these categories were refined by the project staff, they were then sorted into "factors".

The project team also organized an advisory group made up of the users of the system (such as enlisted assignment managers, personnel management policy decision makers, etc.) and command enlisted advisors. The advisory group provided a review of the panel results, and helped to decide if the personal performance characteristics would be useful in everyday decisions that involved the use of enlisted evaluation marks. Decisions on which of those characteristics that should be kept were based on which would help:

(1) identify training needs;

(2) provide feedback to evaluee on job performance;

(3) identify potential for advancement;

(4) make selections for job assignments/transfers;

(5) identify the marginal/submarginal performer for possible separation/discharge from the service or reduction for incompetency;

(6) determine eligibility for Coast Guard Good Conduct Medal;

(7) determine type of discharge/reenlistment.

The final results were streamlined into 51 personal performance characteristics. The characteristics were consolidated on the forms to 29 for non-rated, 38 for petty officers, and 37 for chief petty officers. A majority of the characteristics were used for all three groups. Those characteristics that were the same or had the same meaning were given different performance standards to allow for the different responsibilities between the pay grade groups.

For example, the low, middle, and high standards for a chief petty officer (E-7) were written higher than for a seaman (E-3) because the organization expects a higher caliber of performance from the senior member. This methodology was also designed to promote the "different pay grade, different standards" theory.

Participation and involvement from the Coast Guard population in designing the new system was also advantageous in the implementation phase. Sixty-five facilitators were selected from each of the twelve Coast Guard districts to conduct a month long servicewide orientation and training program on how to use the new system. These "hand-picked" facilitators made up mostly of noncommissioned officers from various vocational backgrounds, knew nothing of the new system, and were selected on their personal and professional abilities and qualities as effective communicators. The facilitators underwent an intensive 30 hour training program on the philosophy and mechanics of the new system. After completing the program, they conducted eight hour orientation and training programs at Coast Guard units in their respective geographic areas.

Summary

The shortcomings of the old system and many of the proposals recommended by the theses and OSU study were addressed in the design of the new system. The system emphasizes a clean break from the old system. A three form concept will separate the pay grade groups, with each form having a different number of characteristics. The pay grades will have separate evaluation periods to stress the differences in experience and responsibility. A seven point numerical scale will be used, with the even numbers assigned the marginal, expected, and exceptional performance standards. These standards as marked, will provide feedback from the supervisor to the subordinate. Although it gives a more extensive evaluation, the new system is only slightly more complex than the old. The guidelines remain simplistic and require only a period of structured orientation for people newly exposed to the system.

The revised system is currently in the monitoring stages, where the project team will be analyzing the fall and spring evaluation periods to determine if the new system has met the organization's expectation, and if the personnel management policy decisions that are tied to the system are sufficiently defined.

References

Stumpff, J. F. and Chevalier, R. D. An analysis and proposal for revision of the Coast Guard enlisted performance evaluation system. Thesis submitted to the Naval Postgraduate School, Monterey, California, December, 1976.

Cornelius, E. T. and Hakel, M. D. A study to develop an improved enlisted performance evaluation system for the U. S. Coast Guard. Final report for contract DOT-CG-62625-A, February, 1978.

Flanagan, J. C. The critical incident technique. Psychological Bulletin, Vol. 51, No. 4 (1954), 327-358.

Shields, W. S. A redesigned performance appraisal system for noncommissioned ranks in the Canadian Armed Forces. Paper presented at the annual meeting of the Military Testing Association, San Antonio, Texas, November, 1982.

Use of Self Assessments in Estimating Levels of Skill Retention

Joel D. Schendel and John C. Morey
US Army Research Institute Field Unit
Fort Benning, Georgia

M. Janell Granier and Sid Hall
Auburn University

Abstract

The focus of this research was on one approach toward predicting task retention and refresher training requirements--self assessment. Prior to zeroing their weapons for annual M16A1 rifle qualification (record fire), 153 permanent party soldiers completed a questionnaire designed to collect information on their previous marksmanship experiences as well as their beliefs about how they were going to shoot at record fire. No special instructions were provided about the meanings of questions or possible responses. Questionnaire items then were correlated with rifle qualification scores based on 40 target exposures. Predicted scores accounted for about 10% of the variance associated with record fire scores. However, for those soldiers whose confidence in the accuracy of their predictions exceeded 90%, predicted scores accounted for 25% of the record fire score variance. The best predictors of record fire performance were remembered most recent record fire performance and predicted performance. Subjects, generally, overestimated their actual performance and were biased heavily toward predicting success. The usefulness of the self-assessment approach to skill retention estimation is discussed together with practical suggestions toward refining the precision of questionnaire techniques.

Introduction

Most Army jobs require that soldiers achieve and maintain proficiency on scores of tasks. Because of resource constraints it is impossible to train every soldier on every task to the degree necessary even to minimize forgetting. As a result, some amount of periodic refresher training must be provided. This much is clear. What is not clear is how frequently individuals require refresher training to sustain proficiency on particular tasks. As Schendel and Hagman (1982) indicate, if time intervals between successive sessions are too long, then performance may fall below acceptable levels and entail considerable risk. Emergencies may arise requiring corrective action before an individual has had a chance to retrain. If time intervals between sessions are too short, then administrative costs necessarily are inflated.

What can be done to remedy this problem? One possible approach toward managing refresher training involves the use of self assessments . Under this approach, priorities would be assigned to tasks to be trained (managerial decision). Unit commanders then would determine where they are weakest, next weakest, etc by sampling soldiers' estimated skill levels on these tasks. Priorities would be weighed against weaknesses as well as resources (time,

travel requirements, equipment constraints, etc) and unit refresher training schedules developed accordingly. Of course, self assessments need not be used alone. They may be used to identify soldiers requiring training on particular tasks and, consequently, may be used to refine predictions derived from other, more generalized approaches (e.g., Rose & Ford, 1982).

The advantages of the self assessment approach are that it is simple and economical--self assessments can be collected using only a questionnaire. Self assessments also can be very accurate (e.g., Levine, Flory, & Ash, 1977). The main disadvantage of self assessments is that they are subject to distortion, both intentional and unintentional. For example, the stereotype of a "good soldier" may influence self assessments. As a result, soldiers may be reluctant to admit needing training on tasks which they should know. Alternatively, soldiers may unintentionally overestimate their abilities to perform a task as a result of their overconfidence. Several methods for handling these problems are discussed later in this report. Many of these methods also have been covered by Burnside (1982). This research examines the feasibility of using self assessments in predicting soldiers' task retention and refresher training requirements.

Method

Subjects

The subjects were 153 male (n = 147) and female (n = 6) permanent party soldiers assigned to the 1st and 2d Infantry Training Brigades and Infantry Training Group at Fort Benning, Ga., who fired record fire between 26 and 27 March and 4 and 5 June 1982, and who volunteered to participate in this research. Subjects completed the questionnaire in groups ranging in size from approximately 15 to 60. Reported times since last record fire ranged from 1 to 60 months (n = 147, M = 12.80, SD = 10.93).

Materials and Procedure

Questionnaire data. Prior to firing record fire, soldiers reported to a 25-meter range to zero their weapons. Zeroing involves firing series of three-round shot groups at a specially designed 25-meter target and adjusting the rifle's sights until point of bullet impact coincides with point of aim. Prior to zeroing, soldiers assembled in bleachers to receive a safety briefing. After soldiers had assembled in the bleachers but before they had received this briefing, they were informed about the purpose of this research. Subjects were told merely to answer the questions as accurately as possible. No special instructions were provided about the meanings of questions nor were possible responses discussed.

Performance data. After zeroing, all soldiers were transported to an M16A1 rifle qualification range. On arriving, each subject was assigned randomly to one of eight firing lanes. All scoring was done by independent support personnel. These personnel were fully informed about scoring procedures and the purpose of this research prior to the onset of testing. In addition, an experimenter and numerous range personnel were available to assist in scorekeeping and to answer questions.

The course-of-fire consisted of 14 "F-type" silhouettes and 26 "E-type" silhouettes. F-type silhouettes are designed to appear like the head and shoulders of a man and were seen at 50 (n = 5) and 100 (n = 9) meters. E-type silhouettes, designed to appear like the head and torso of a man, appeared at 150 (n = 10), 200 (n = 8), 250 (n = 5), and 300 (n = 3) meters. The first half of this course was shot from the foxhole supported position; the second half from the prone unsupported position. Targets fell when hit. Record fire scores and related shooting classifications are as follows: 0 to 22--Unqualified; 23 to 29--Marksman; 30 to 35--Sharpshooter; 36 to 40--Expert.

Results

Pearson correlation coefficients (rs) first were computed on all subjects' questionnaire and record fire data. Subjects' predicted and actual record fire scores correlated positively, r = .26 (n = 150, $p < .01$). A similar result was obtained when computing the r between subjects' predicted shooting classifications and their actual classifications, r = .25 (n = 152, $p < .01$). While these rs appear low, they were much stronger than others which we expected may be quite strong. For example, reported experience outside the Army firing a rifle or a shotgun failed to correlate with performance at record fire, r = −.10 (n = 152, $p > .05$).

One of the best predictors of record fire performance was soldiers' remembered shooting classifications from their most recent record fire, r = .38 (n = 133), $p < .01$). Remembered most recent record fire scores also correlated with record fire performance, r = .29 (n = 121, $p < .01$). Also, remembered classifications and scores related highly to those predicted. Remembered classifications correlated r = .65 (n = 132, $p < .01$) with predicted classifications and r = .53 (n = 130, $p < .01$) with predicted scores; remembered scores correlated r = .67 (n = 120, $p < .01$) with predicted classifications and r = .63 (n = 118, $p < .01$) with predicted scores.

More extensive analyses then were carried out on questionnaire data corrected for internal consistency. This correction involved determining whether each subject's response to the question "How many targets out of 40 do you think you will hit at record fire?" was consistent with his (or her) responses to the following questions: "What is the highest number of targets out of 40 you feel you are likely to hit today?" "What is the lowest number of targets out of 40 you feel you are likely to hit today? For example, if a subject indicated he (she) believed he (she) was going to hit 27 targets at record fire, but then indicated that the highest (lowest) number of targets he (she) felt likely to hit was less (more) than 27, the subject's data were not included in the analyses that follow.

Most of the rs computed following this correction were similar in magnitude and direction to the rs reported above. Particularly noteworthy, however, was the r between soldiers' predicted and actual record fire scores which showed some improvement, r = .32 (n = 124, $p <.01$). In addition, the r between reported months since last record fire and record fire performance achieved significance, r = −.19 (n = 123, $p < .05$). Reported months since last record fire should correlate negatively with record fire performance, at least to the degree that this variable reflects the length of the interval between successive refresher training periods.

As indicated earlier, subjects were asked to predict their shooting classifications. They also were asked to estimate their chances of being correct in these predictions. In general, subjects who estimated their chances of being correct at 90% or 100% were more accurate in their predictions than subjects who gave lower confidence estimates. In particular, predicted shooting classifications and record fire scores correlated $r = -.01$ ($n = 36$, $p > .05$) for subjects who estimated their chances of accurate prediction at less than 60%. This r was .24 ($n = 47$, $.05 < p < .10$) for subjects who estimated their chances of accurate prediction as being between 60% and 80%. And, r equalled .42 ($n = 41$, $p < .01$) for subjects who estimated the likelihood of accurate prediction at 90% or 100%. Similarly, the rs between predicted and actual record fire scores were nonsignificant for subjects who estimated their chances of accurate prediction at less than 60%, $r = .23$ ($n = 36$, $p > .05$) or between 60% and 80%, $r = .04$ ($n = 47$, $p > .05$). However, this r was highly significant for subjects who estimated their chances of accurate prediction at 90% or 100%, $r = .50$ ($n = 41$, $p < .01$).

Estimation errors (i.e., algebraic errors) were inversely related to record fire scores, $r = -.73$ ($n = 124$, $p < .01$). An examination of these errors indicated that the predictions of subjects who fired well at record fire, generally, were more accurate than those who fired poorly. This suggests that subjects who performed well also were more skilled at self assessment. On the other hand, this result may merely reflect a general tendency to overestimate shooting ability: 75% of the subjects predicted they would hit more targets than they actually hit. If most subjects thought they would do well, good shooters naturally would show less error in their predictions than poor shooters.

Only five (4%) subjects predicted that they were going to fail to qualify (i.e., hit less than 23 targets). Actually, 34 (27%) subjects failed. More interestingly, however, three of the five subjects who predicted that they were going to fail did in fact fail. And, the remaining two who passed, passed only by two points. Conclusions are limited by the small sample size. However, this result suggests that, while subjects were biased heavily toward predicting success, those who predicted failure were quite accurate.

Discussion

Correlations between predicted and actual scores were not high. At best, predictions accounted for only about 10% of the variance associated with record fire scores. Nevertheless, these data must be regarded as encouraging for the following reasons:

Record fire scores are notoriously unreliable. Equipment failure, scorer bias, variations in light conditions, vegetation, and terrain all contribute to this unreliability (e.g., Marcus & Hughes, 1979).

No special instructions were provided subjects as they completed the questionnaire. The idea was to obtain a baseline estimate of soldiers abilities to assess their own skills under field conditions.

The data appear logically consistent. Predicted and actual performances correlated positively as did predicted and remembered performances. The correlation between predicted and actual scores showed some improvement when data were corrected for internal consistency. And, subjects who expressed the most confidence in their predictions also were most accurate.

Future research should focus on the problem of improving the accuracy of subjects' self assessments. Several manipulations already are known to improve this accuracy (e.g., Burnside, 1982). Our experience conducting this research suggests at least three further manipulations:

1. Provide subjects as much information relevant to the formation of accurate self assessments as possible. This information is conveyed largely by descriptions of the task, conditions, and standards. However, it also may be beneficial to remind subjects how they did previously or to refresh their memories for a task using a demonstration.

2. Have subjects provide confidence ratings along with their self assessments. Confidence ratings appear a good index of the accuracy of these assessments. In this research, the predicted scores of subjects who estimated their chances of accurate prediction at 90% or 100% accounted for 25% of the variance associated with record fire scores. Similar results have been obtained by Fischhoff and MacGregor (1981).

3. Instruct subjects to check their responses for internal consistency as they complete their self-assessment questionnaires.

References

Burnside, B. L. Subjective appraisal as a feedback tool (ARI Technical Report). Alexandria, Va: U.S. Army Research Institute for the Behavioral and Social Sciences, May 1982.

Fischhoff, B., & Mac Gregor, D. Subjective confidence in forecasts (PTR Technical Report PTR-1092-81-12). Woodland Hills, Ca: Perceptronics, December 1981.

Levine, E. L., Flory, A., & Ash, R. A. Self-assessment in personnel selection. Journal of Applied Psychology, 1977, 62, 428-435.

Marcus, A., & Hughes, C. R. An evaluation of a technique for using combat training theater (CTT) for periodic rifle marksmanship proficiency training and qualification (ARI Research Problem Review 79-7). Alexandria, Va: U.S. Army Research Institute for the Behavioral and Social Sciences, May 1979.

Rose, A. M., & Ford, P. Acquisition and retention of soldiering skills: Report of year 1 progress (AIR Technical Report AIR/IR-88601-4/82). Washington, DC: American Institutes for Research, April 1982.

Schendel, J. D., & Hagman, J. D. On sustaining procedural skills over a prolonged retention interval. Journal of Applied Psychology, 1982, 67, 605-610.

Comparing Two AFOQT Norming Groups Using Weighted Multidimensional Scaling

Bennie W. Roach
Air Force Human Resources Laboratory
Brooks AFB, Texas 78235

Abstract

In test development, finding a sample of subjects that is both accessible and representative of the target group is often a difficult process. The Air Force Officer Qualifying Test (AFOQT) is administered to civilian college students and graduates applying for precommissioning programs. If Basic Airmen undergoing training at Lackland AFB are a sample of the civilian applicant population, they can be used as a convenient and inexpensive source of research subjects for development of new items and subtests for the AFOQT. This study used the individual differences option of the PROC ALSCAL multidimensional scaling procedure to compare a sample of Basic Airmen with a sample of actual AFOQT examinees. AFOQT subtest intercorrelations were used as input stimulus data. Analysis of subject spaces representing the two samples indicates that there is essentially no variance and, therefore, the two norming groups are from a single target population.

Introduction

The Air Force Officer Qualifying Test (AFOQT) is a test battery used as an aid in the selection of applicants for precommissioning training. The test is also used to select and classify new commissionees into pilot and navigator training.

Applicants for precommissioning training in the Air Force are limited to college students (Air Force Academy and Air Force Reserve Officer Training Corps) and college graduates (Officer Training School). However, the norming and standardization of new test forms has traditionally used enlisted high school graduates who are undergoing Basic Military Training (Basic Airmen) since they are readily available for large-scale testing. Both samples are administered the AFOQT. The Basic Airmen are administered the test involuntarily for research purposes. The actual AFOQT examinees are administered the test on their request for the purpose of applying for one of the precommissioning programs in the Air Force.

Although never verified, it has been generally assumed that Basic Airmen (who are high school graduates and mental Category I & II) and actual, or "true," applicants for precommissioning are essentially two

samples of the same population. If this is true, Basic Airmen can be used for field testing of new test items, subscales, and forms. The distribution of performance by Basic Airmen can then be adjusted for ability between the two samples in order to develop new conversion tables. However, if Basic Airmen and actual applicants are from two different populations, no amount of psychometric adjustment will provide accurate conversion tables.

The problem investigated by this study can be stated as follows:

> Is the group traditionally used as a normative base in the development of the AFOQT test battery a sample of the same population which includes actual applicants for Air Force precommissioning training?

This problem can be restated as the following hypothesis:

> Basic Airmen and actual AFOQT examinees are two samples of the same population.

Methodology

The AFOQT is a test battery composed of the following 16 subscales: Verbal Analogy, Arithmetic Reasoning, Reading Comprehension, Data Interpretation, Word Knowledge, Math Knowledge, Mechanical Comprehension, Electrical Maze, Scale Reading, Instrument Comprehension, Block Counting, Table Reading, Aviation Information, Rotated Blocks, General Science, and Hidden Figures. These 16 subscales were used as the input variables. Although each subscale of the AFOQT is designed to test a unique facet of aptitude for training, there is a certain amount of similarity between the subscales. The product-moment correlation coefficient was used as the measure of similarity. In multidimensional scaling terms, each subscale served as a single "stimulus," and the correlation between pairs of subscales (paired stimuli) served as indicators of "proximity." For each of the two samples, the lower, off-diagonal (120 elements) of the 16 x 16 matrix of pair-wise proximities was used as the input to an individual differences (weighted) multidimensional scaling model.

Subjects

The most recent version of the AFOQT, Form O, was administered to a first sample (S_1) composed of 952 Basic Airmen in an experimental setting. The test battery was also administered to 17,059 actual applicants for precommissioning training. This second sample (S_2) took the test at Air Force bases and Military Entrance Processing Stations throughout the United States.

Analysis

Since the purpose of this study was to determine if the two samples of AFOQT examinees came from the same population, it was necessary to separate information that was common to both samples from information that was unique to each sample. This was accomplished by treating the two sets of input proximity data as representing two single subjects. The PROC ALSCAL multidimensional scaling software contained in the Statistical Analysis System (SAS) was used to transform the two sets of proximity data into distances, to configure these distances in a multidimensional space, and to compute individual subject weights on each dimension. The dimensional differences between the two subjects were determined by means of the PROC ALSCAL individual differences option. The criterion for the appropriate number of dimensions required for the solution is interpretability, with goodness of fit, as measured by "stress" and "squared correlation," being used as aids in making the final determination.

Results

PROC ALSCAL was performed on the proximity data for two-, three-, and four-dimensions. Solutions of higher dimensionality were not used because of difficulty in interpretation. The results of the measures of goodness of fit are provided as Table 1. The amount of stress for each matrix and the average of the two matrices were reduced with the addition of each new dimension, while the amount of squared correlation increased. Thus, of the three solutions, the one using four dimensions provided the best fit.

Table 1

Measures of Goodness of Fit

Number of Dimensions	Matrix	Stress	Squared Correlation
2	S_1	.213	.748
	S_2	.196	.768
	Average	.205	.758
3	S_1	.145	.832
	S_2	.132	.853
	Average	.139	.843
4	S_1	.106	.888
	S_2	.085	.929
	Average	.096	.908

The individual differences option on PROC ALSCAL provides a plot of the group stimulus space for each two-dimensional combination and a corresponding plot of the subject space for each combination. Due to limitations on the length of this paper, the plots were not included. The subject space provides a spatial representation of the weights for each subject (matrix). Table 2 provides the respective subject weights for each of the three solutions.

Table 2

Subject Weights

Number of Dimensions	Subject*	Dimension 1	2	3	4
2	1	.5648	.6551		
	2	.6846	.5473		
	Average	.6247	.6012		
3	1	.6549	.5602	.2993	
	2	.6730	.5674	.2791	
	Average	.6640	.5638	.2892	
4	1	.6450	.4072	.4118	.3692
	2	.5744	.5154	.4790	.3226
	Average	.6097	.4613	.4454	.3459

*Subject 1 = Basic Airmen
*Subject 2 = AFOQT Examinees

The four-dimensional solution provided two-dimensional plots for the group stimulus spaces and subject spaces for each of the six possible pairings of the four dimensions. The group stimulus space for Dimension 1 versus Dimension 2 provided the most perceptual meaning. This two-dimensional plot displayed the subscales in clusters which represented the composites to which they belong.

While the four-dimensional solution is interpreted as the optimum solution, the initial problem has yet to be answered, e.g., are the two samples from the same population. The answer is obtained by performing a statistical analysis of the subject spaces obtained from the four-dimensional solution. The subject weights were normalized and the statistics computed for the total solution and for each of the two-dimensional spaces. The mean resultant length and the index of angular variation for each of the six two-dimensional pairings are provided in Table 3.

Table 3

Subject Space Statistics
(Solution in Four Dimensions)

Horizontal Dimension	Vertical Dimension	Mean Resultant Length	Index of Angular Variation
1	2	.9965	.0035
1	3	.9980	.0020
1	4	.99999	.00001
2	3	.9998	.0002
2	4	.9961	.0039
3	4	.9976	.0024

The index of angular variation uses a 0 to 1 scale to indicate the degree of clustering of the subject vectors based on normalized subject weights. Thus, when the index is zero there is no variation and the subjects are identical. The total angular variance represented by the index is separable into "within" and "between" components (similar to ANOVA) with a procedure called Analysis of Angular Variation (ANAVA). As Table 3 shows, none of the index values were greater than .004. Therefore, there is no statistically significant difference between the two samples, and the null hypothesis is accepted that the two subjects (samples) are from the same population.

Conclusions

Two conclusions are supported by the results of this study: (1) for purposes of AFOQT testing, Basic Airmen (Cat I and II) and actual AFOQT examinees can be considered as two samples from a single population; (2) multidimensional scaling models provide a more in-depth approach to comparing samples than the more traditional procedures, which do not evaluate differences in the dimensional composition of samples.

TRAINING DECISION-MAKERS TO BE CREATIVE: A MANAGEMENT PROCESS MODEL

Beth Ann Martin
Kathryn Strom-Guzowski
Robert L. Taylor

Abstract

The predictable structure of our military bueaucracies reinforce traditional decision-making processes. At the same time, complex technologies and environments suggest more creative decisions. This research focuses on how experienced decision-makers can be trained to be more creative.

Introduction

Our purpose is twofold. First we examine how decisions are made. Examining models of decision making may help us to better understand the nature of what we do and where creativity fits in. Second our task is to study the research in creativity, developing a model that relates it to decision making.

The major focus is integrating the two models--decision-making and creativity--in a theoretical framework. This is a first step; before a training program can be developed, the theoretical roots must be verified.

Method

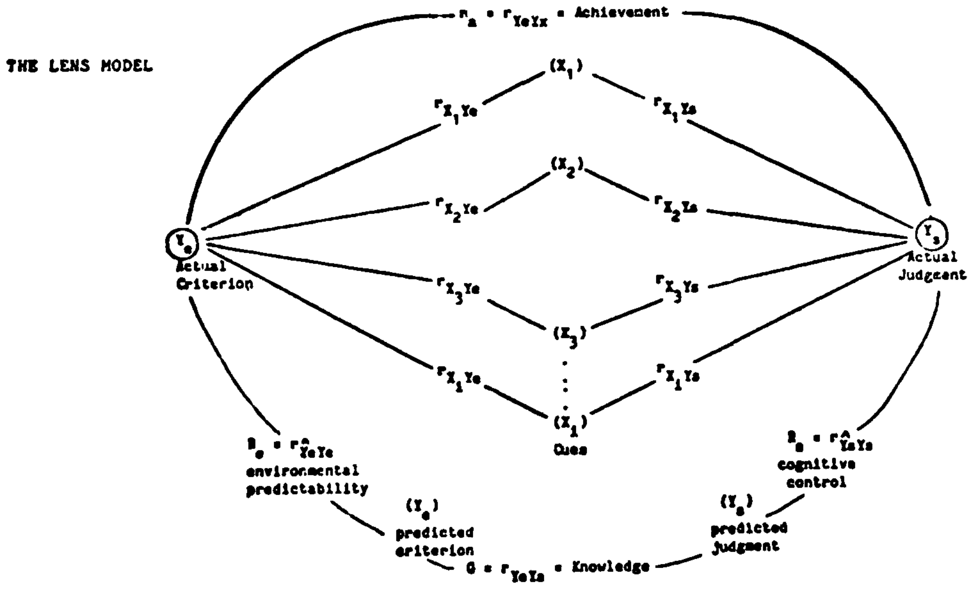
One decision-making theory is described. Second, an explanation of the creative process is given. Third, we integrate the two models. Fourth, we discuss the implications for encouraging and training people to be more creative decision-makers.

Decision-Making. Decision theory is a complex body of literature which attempts to prescribe how decisions should be made as well as defining the variables that systematically affect our decisions. Our focus as we begin, is with events occuring at the cognitive level, understanding the processes and strategies used in integrating discrete units of information into decisions.

In the 1960's and early 1970's researchers found linear models could describe cue weighting policies in a laboratory. Research then moved into real-world judgments where business managers, military officers, accountants, Air Force Cadets and theatrical plays were studied. This approach has been able to account for most of the predicted variance in specific complex judgments made by individuals within these respective areas. It has also provided useful descriptions of the judges' cue weighting policies--pinpointing the sources of interjudge disagreement. Because of the practical utility of the regression approach, a specific regression theory, Social Judgment Theory (Hammond, 1975) is selected for use.

Social Judgment Theory uses correlation statistics to provide judgmental models in realistic settings. It suggests individuals must necessarily learn to function and to adapt to

situations in which the information they have available is not a completely reliable indication of social reality. It is the task of the individual to identify the most useful sources of information and combine them in order to maximize one's adaptation to the social environment. The individual must identify the most important cues, using those in making effective judgments. The lens model diagram (Brunswick, 1956) best portrays Social Judgment Theory. The lens model may be viewed as any situation in which an individual is making judgments. The individual's task is to identify the most useful sources of information and learn to employ them skillfully.



Ye = Actual Criterion.
X1 . . . Xi = cues available to judge.
Ys = Actual Judgement.
ra = Achievement: the judge's success in predicting over a number of trials.
rXiYe = cue validity (ecological validity).
rXiYs = Cue Utilization: an index reflecting the extent to which a judge is depending upon a particular item of information.
Rs = rYsYs = Cognitive control: this is the degree of consistency a judge employs from time to time based on one set of cues.
G = Knowledge: a measure of the extent to which the individual has correctly detected the properties of the task.

The focus here is on capturing the cognitive policy of the individual. That is, identify the cues and weighting schemes used by the individual decision-maker (a judge).

We've isolated Social Judgment Theory for our purposes because of its wide applicability for all types of judgments. Its major strengths for our purposes are the theory's ability to indicate what cues people use when making decisions, plus the level of importance people attach to these cues. Knowing an individual's cues then allows one to provide feedback to individuals regarding how they are using those cues when making personal decisions. This feedback can serve a variety of

functions one of which is as the basis for a training program for improving efficiency in making decisions.

Creativity. Within the field of creativity, there is much information which remains a mystery. One area of study relates to the creative process, or the specific problem solving methods used by individuals who are considered creative. The identification of the cognitive components of creativity are best described in terms of the components of the problem solving process.

A five-step model proposed by Dewey (1910) has been selected for use because its terms can be easily operationalized. This model includes recognition, analysis, generation of solutions, testing and judgment. Our emphasis is on recognition and analysis, which encompass problem formulation, and the generation of solutions. In these three stages there is a direct relationship between creativity and problem-solving.

Recognition: the perception of past and present environmental experiences which aid in defining the specific nature of a problem. In this stage, an individual demonstrates a sensitivity to problems and then systematically examines the environment to determine the source of the problem.

Analysis: the discovery of the location, nature, and meaning of the problem based upon the perceived experiences. Synthesizing the environmental cues leads to a clear statement of the problem.

In specific instances, recognition and analysis may be combined into a single stage. This would be the case when one is presented with a problem and must then come up with a solution to the problem. In this instance, no creativity is required to formulate the problem.

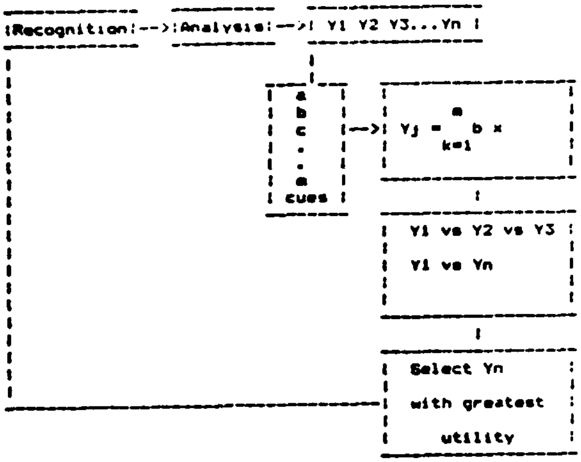
Generation of Solutions: involves the development of any number of potential solutions for a problem. Fluency and flexibility of thought are critical to the generation of solutions. The creative person will develop more potential solutions than the less creative person. The quality of the solutions also distinguishes the creative from the less creative individual. Potential solutions must be relevant, practical and unique.

Testing: the process of comparing a solution with every other solution so that the best solution can be determined.

Judgment: the selection of one solution from all those possible solutions; the solution with the greatest overall value.

Integration

A model of the integration of decision-making and creativity provides a useful starting point in the effort to show that creativity and decision-making involve two aspects of the same process, problem-solving. Hopefully, by gaining a fuller understanding of this process, we can begin to maximize our decision-making potential.



The first step is that of problem recognition. One could delineate two distinct problem types. The first (Type I) would involve an individual's sensitivity to the environment such that an existing problem is recognized. The more creative person while existing within his/her environment, notices something amiss and self defines a problem. In a Type II situation the problem has been previously defined. The person is aware of the problem to which the decision should apply and initiates an analysis of the problem at that time. Thus, one need not deal solely with an already defined problem and thus limit oneself to measuring the creative product. Being open to one's surroundings and recognizing and defining a problem is also part of the creative process.

Step 2 is the analysis of the problem. Here one uses information offered by the immediate environment, searches the environment for information or retrieves information from memory. The available information must be utilized to determine the exact nature of the problem before attempting to generate alternative solutions. At this point, the problem should be dissected into its smallest elements; breaking down what we know and tearing apart the structures which cause our generally routinized decision-making.

Once the problem has been torn down a building up process can begin. By putting the elements presently recognized and those from past experience together in new and different ways alternative solutions are generated (Step 3). A creative individual will rearrange these elements many times in order to generate a variety of solutions. The non-creative person will simply do a cursory job of problem analysis and use only past and existing information in old, non-novel combinations to generate a very few redundant solutions. In essence, the creative person will show more fluency in thought during the generation of solutions, resulting in the production of more possible solutions than the less creative individual.

After a number of potential solutions are generated the individual will determine a variety of cues, based upon past experience and relevancy to the problem, to evaluate the now available alternatives (step 4). In a creative person, past cues will no longer be salient, hence new cues will be generated. In the less creative person the past cues are selected.

At step 5 each solution is tested according to many combinations of the proposed weights. Flexibility permits one to

rearrange the cues in terms of their significance to the problem and assign them a variety of weights. Thus, each solution may be pondered and weights assigned in a number of ways as various conditions are reweighted.

At step 6 the weighting of the cues for each solution is complete. A comparison is made among the alternatives. Finally at step 7, the alternative with the highest utility is decided upon. Thus, we have a model which indicates the creative process impacting on decision-making at three major points: when an individual notes, recognizes and interprets a felt difficulty as a problem; in the novelty of solutions generated; and how the individual uses the cues. Are many or only a few cues generated? Does the individual weight and reweight solutions when initially trying to use these new cues? These are key indicators.

Social Judgment Theory allows us to break down overt judgments and infer the cognitive processes which occur. An additional concept, creativity, can be seen as a concurrent problem solving process. An examination of these two processes suggests they are not distinct parallel cognitive processes.

Conclusions

Everyone has the ability to be creative. The thing that seems to inhibit the development of creative potential is the environment in which we live and work.

If we really want people to be "more creative" in their decision-making, we must reward the product that results. Simply by asking people to take another look at an issue, consider unfamiliar alternatives, or start over again, we give the signal that another alternative might be appropriate.

We can ask for several solutions to a problem. Then, by asking questions (being interested in some of the more unlikely solutions), a message is sent that the tried and true response might not be the one most desired. By casting about for multiple solutions, we automatically extend the range of interest, thereby forcing people to consider things they wouldn't have thought of before.

Encouraging creativity is time-consuming. Thus, we don't want to change behaviors relative to <u>all</u> decisions. The focus can be limited--with specified problems or opportunities we expect more creative products to evolve. In essence, we develop two classes of decisions: 1) the routine, patterned responses, and 2) the non-routine, creative responses. There are different <u>processes</u> for each.

References

Brunswick, E. <u>Perception and the Representative Design of Psychological Experiments</u> (2nd ed.). Berkeley: University of California Press, 1965.

Dewey, J. <u>How We Think</u>. New York: D. C. Heather, 1910.

Hammond, K. R. et. al., Social Judgement Theory. In Kaplan and Schwartz (eds.). <u>Human Judgement and Decision Processes</u>. New York: Academic Press, 1975.

On the Meaning of the "Whole Person" Concept and Achieving Wholeness

Charles Jacob Yoos II
United States Air Force Academy

Abstract

Production of a "whole person" is a key feature in the architecture of many professional military development programs, including that of the Air Force Academy, but the concept of "wholeness" is widely misunderstood to mean *variety* of knowledge and experiences rather than *integration* of them. In this paper, I define what wholeness really means, and explain why it is so easily misconstrued. Then, I offer four methods of achieving wholeness: archetypal emulation, metaphors, mysticism, and metatheory, and argue a preference for metatheory.

The "whole man concept" (I will use the more contemporary, genderless "whole person concept") has been a fixture in the Air Force Academy's ideology, but it is generally misunderstood. Most officers who espouse it convey the perception that a whole person is produced by causing that person (here, the cadet) to be exposed to, experience, and learn a *great variety* of things, and the greater that variety, the "more whole" the person will be. Some also perceive the need for juxtaposition, or logical sequences, of certain of those things.

These perceptions miss the mark! It is critical that cadets be given a broad and rich regimen in some systematic fashion, but that alone does not render them whole any more than the Space Shuttle assumes identity merely by getting together, in order, x tons of titanium, y miles of cabling, and z numbers of ablative tiles--it is obvious that the integrity (which, by the way, means wholeness) of the Space Shuttle depends on the appropriate configuration of the pieces, so that they all "fit" together. What seems not to be realized by most is that the analogy holds: no person can be said to be "whole" unless all the pieces of their knowledge and experience are *coherent*. A person with less, but better integrated knowledge is actually more of a whole person than one with more, but not well put together, information. Wholeness cannot be arrayed on the customary dimensions of intensity and extensity, or even sequence--it is orthogonal to them. This point must be given explicit attention in our stewardship of professional military development programs.

At the Academy, the faculty officer who decides that intellectual wholeness is served by an interdisciplinary course offering, and approaches it as a "you take so many lessons and I'll take so many lessons" proposition, falls prey to the wholeness fallacy. The Commandant's staff officer who strives to produce whole 2nd Lieutenants by assembling the body of knowledge that "every 2nd Lieutenant needs to know" conducts a futile search. The athletic instructor who tries to arrange bodily wholeness by exposing the cadet to a

host of basic and "carry-over" sports, commits the same error. In each case, they mistake *accumulation* for *conglomeration*.

Why is the essence of wholeness so easily misconstrued? Here are several interrelated reasons:

> -- It is a difficult concept to operationalize. On the other hand, if it is (mis)understood to mean nothing beyond exposure to more and different things, then that is more easily implemented, so it is natural that a mind grappling with wholeness is prone to this error.
>
> -- The process is outside our locus of control as instructors, commanders, coaches, etc.--it can't be engineered, and no algorithm describes it. Like the sprouting of a seed, it can be nurtured, but it occurs on its own.
>
> -- The process of becoming whole is time consuming, and hence expensive. What is more, there are no tangible indicants along the way (like graded recitations) to reassure of progress. The time required for it can't be scheduled, and some of that time even appears to be idle.

For all of these reasons taken together, the mind, which strives for cognitive consistency within a means-end framework, tends to block out an interpretation (however accurate) with which it cannot, or prefers not, to cope.

How <u>can</u> wholeness be achieved? One classic strategy is to create an *archetype,* sculpted as a paradigm of wholeness, and then enforce compliance. This does not guarantee success, but in a relatively closed system like the Academy, if the initiate is diligent and sincere, then the chances of becoming whole through archetypal emulation are enhanced. One variant on this theme is to incarnate the archetype by emulating an actual appropriate "great person" who is seen as whole. A child who emulates a parent is an ordinary but not trivial example. If, as seems likely in the professional military situation, the great person is historical, then getting information and transferring contexts can be problematic--suppose that I decide to be like General George Catlett Marshall (1880-1959): can I learn enough about him to emulate successfully; even if I can, will I be an anachronism? This problem is obviated, of course, by emulating a contemporary--expanding the concept of a "role model" to that of a "whole model," but the fact that wholeness is usually ratified by time completes the dilemma.

If the archetype remains disembodied, then a comprehensive codification is typical. The image here is of medieval knights, becoming whole through chivalrous conduct as pages and squires. The crux of any such approach is to fully delineate and sanction "what it means to be us." Aside from the inherent problems of constructing and administering a code robust enough to result in wholeness, there are two other troubles with this alternative for a military service academy in our society:

> -- Such a paradigm, which usually incorporates lofty standards of conduct, lends itself to popular manifestation as a caste system, where "we" are perceived by others to consider ourselves to be

better than "they," and this is anathema to a society premised on class equality. The paradox is that taxpayers have every right to expect a superior product for their investment, but it must be packaged advisedly in this regard.

-- A single archetype for emulation denies the diversity of excellence that an institution like the Air Force Academy can provide to the Air Force and society, where graduates are expected to be superior, but not uniformly so.

A second method of achieving wholeness is to invoke *metaphors*. The power of a metaphor is not isomorphism; rather, it is in finding striking similarity within stark difference. Here, the search is for a seemingly unrelated domain of meaning where "wholeness" is vivid. Earlier in this paper, for example, I asserted a similarity between a whole person and the Space Shuttle, a metaphor designed to confirm the distinction between the assemblage and the integrity of parts applied to people as well as machines, but not intended to convince you that a whole person might blast off and orbit the earth! Another metaphor might be to think of a whole person like a symphony orchestra, not that one's esophagus is like a bassoon, but that whole thought is harmonious.

Two aspects of metaphors bear emphasis. First, it probably works best to entertain a richness of many metaphors of wholeness, allowing them to converge in the mind. A "master metaphor" of wholeness is unlikely. Second, my years of using metaphors with undergraduate students in organization theory have taught me that a metaphor can be offered to, but never imposed on, an inquiring mind; indeed, the most potent metaphors spring from within. Thus, making metaphors for wholeness (or anything else) ought to be presented as exploratory technique, not particular solution.

A third means for wholeness, though one which might well be viewed as unacceptable from our occidental cultural perspective, is *mysticism*. Here, through instruction followed by contemplation, meditation, and ultimately inspiration, holistic wisdom is thought to be achieved. Exactly <u>what</u> is instructed, contemplated and meditated varies over several strains of mystical tradition and is beyond the scope of this paper, but a common theme is the dialectic which reconciles opposites, which finds unity in diversity.

It is unfortunate that the mere possibility of wholeness through mysticism would likely be dismissed out-of-hand and ridiculed by most people, who are probably ignorant that the most staid of the classic sciences, namely physics, is currently entering the throes of a paradigmatic revolution, led by the Belgian Nobel Laureate Ilya Prigogine (1980), the atomic physicist David Bohm (1980), the neurosurgeon Karl Pribram (1971), and others, which promises to reconcile the atomistic and the mystical-wholistic views of the world. In short, these eminent scientists believe that there "is something" to mysticism, and that mystics may well have beat the scientists to knowledge-in-wholeness.

A fourth method for becoming a whole person is to adopt a *metatheory*, a viewpoint which serves to integrate the varied accrued knowledge and experience. This is more than a weltanschauung, in that it must be capable of translating and reconciling <u>all</u> inputs.

The *general systems model* is a metatheory appropriate to wholeness--it provides a conceptual overlay onto which all other ideas may be mapped. The systems metatheory holds that it is advantageous to construe the world as a set of recursive systems, where the behavior of any system cannot be understood by examining the behavior of its parts in isolation. Ancillary ideas like synergy (the output of the whole is greater than the sum of the outputs of the parts), entropy (the inexorable tendency of systems to a state of disorder), and equifinality (a system may achieve a given state via many different inputs and processes) flesh out the systems skeleton.

As with mysticism, several strains of knowledge have developed to apply the fundamental notion of system. Cybernetics treats systems as adaptive mechanisms, responsive to control through feedback. Natural systems theory makes the analogy of social organization to organism. Gestaltism seeks the understanding of configuration totalities, surely familiar to a psychology symposium audience. Cosmogenetic systems theory proposes that the systems comprising the universe are necessarily evolving, responding to perturbation by transcending to higher order existence. Each of these and other strains, while germinating from the common conceptual seed called system, has blossomed in a different (though not contradictory) way, with its own postulates and methods, but what is important to recognize here is that wholeness is the gist of the metatheory.

Metatheory as a method for achieving wholeness is my intellectual preference, for three reasons:

1. It is less troublesome than the other methods. Unlike archetypal emulation, it does not require identification of a relevant historical or contemporary other person, nor does it require comprehensive codification of rules. Unlike metaphors, it does not depend on the creativity of the seeker. Unlike mysticism, it does not subject the seeker to the alien regimen of mysticism. It does, however, require the seeker to embrace the metatheory.

2. By definition, it reconciles the other methods, providing a "wholeness of wholeness methods." Operating from the metatheoretical plane of thought provided by systems thought, the seeker can relate the archetype as an equilibrium system to the tao of mysticism, for example.

3. Perhaps most compelling, it is adaptive. As the world about the seeker changes, and new knowledge and experiences are had, then archetypes may become obsolete, metaphors may become trite, and the mystic tradition is immutable, but a robust metatheory can accommodate the changed circumstances and sustain the vision of the whole.

Wholeness is a legitimate objective, perhaps even the *sine qua non*, of professional military development. No raw amount of knowledge and experience, however, even sequenced intelligently, will produce a whole person. Only the recognition that becoming whole demands integration, not collection, and the use of a valid method, will suffice.

References

Bohm, David. *Wholeness and the Implicate Order*. London: Routledge & Kegan Paul, 1980.

Pribram, Karl H. *Language of the Brain*. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Prigogine, Ilya. *From Being to Becoming*. San Francisco: W. H. Freeman and Company, 1980.

Behavioral Issues in the Management of Technology

Vince Luchsinger, Major General, USAFR
University of Baltimore, Baltimore, MD

Abstract

Considerable discussion today reflects pessimism concerning the effects of technology, and the ability of man to control or manage technology. This paper presents technology as a human activity, capable of being managed. This capability is subject to better understanding and use of human behavior factors involved.

Introduction

The face of technology is usually painted with a mechanistic mien, and pessimistic predictions are made that technology is ruling man in contemporary organizations. This is especially heard in defense organizations and related industrial settings. This paper will look at technology from the management view, and examine some options for improving man's use of technology. Those options will be based on the findings and practice of behavioral science.

Technology and Management. Technology, according to Peter Drucker in his celebrated book, Technology Management and Society (1970) is know-how. This know-how consists of the means that man has devised to help man operate upon his environment, for his purposes. This know-how can be material or non-material.

Material technology comprises devices and machines built with the applications of science and materials that provide extra means for man. This includes aircraft, computers, digital clocks, and ball-point pens among other items. The business of warfare has been an active developer and exploiter of material technology to improve an advantage over potential adversaries. The rise of "high" technology has dealt with rigorous and sophisticated developments of radars, smart weapons, space-craft, and other devices of advanced nature. Lower technologies also play a role in the guise of improved helmets, combat rations, arctic and desert combat clothing. Material technology constitutes a dimension extending from high to low which extends and improves the ability of man to project power.

Non-material technology consists of knowledge, practices, methods, and other non-tangible means of improving man's capability. This includes training, strategy, tactics, logistics techniques, and other means of conducting organizational affairs toward objectives.

In the world of business and industry, that technology (or know-how) is embodied in people, plant facilities, laboratories, and equipment. Effective use of appropriate technology results in a manufactured process, a product, or a service that, consider-

ed as a resource, can be managed to provide a competitive edge in the marketplace. A similar case exists for the military use of technology in the projection of power.

Technology just doesn't happen. Whether as process or product, technology is contrived to provide new options and capabilities for man in pursuing goals. Technology is a means, rather than an end. As a means, technology extends and improves upon man's capabilities.

Some recent reports on technology by the Booz-Allen and Hamilton organization (1981,1982) deal with the provocative issue of technology management. They indicate the key elements are:

1. developing the "right" technology strategy,
2. integrating it with the business strategy, and
3. ultimately making it a part of the total business strategy.

The reports further cite the need to enhance the role of technology in business by greater top-management commitment to longer term results, improved long-range planning of technology, higher R&D expenditures, greater emphasis on integrating technology, and a higher coupling of technology and strategic planning. The reports are not behavioral in nature, but chastise the behavior of decision-makers in business. We can note that decisions and strategies are generated by humans, a crucial thesis of this paper. Technology-related thinking and organization behavior are very relevant topics.

Technology and Man. The history of technology accompanies the history of man. In the Stone Age man worked with primitive tools to deal with his environment, and in this Machine Age, man is using machines. Technology is the product of technical activity by man. Our future will be determined by how we deal with technology. This challenge is affected by the difficulty with which we can see the pattern of our technological environment. Cases of myopia and tunnel-vision abound.

The popular issue of technology transfer is such a case. Technology taken out of the cultural context that produced it is often misused and underused. Machines alone don't make a society technological, a finding seen today in developing countries.

Invention (new processes or products) and innovation (applications) were described by Schon (1967) as the cornerstones of technological progress. Note that Americans developed powered flight, but military applications (innovation) were made on the European continent in WWI while military aviation languished in this country. Application depends upon factors such as public acceptance, economic incentives, and a developmental climate.

Technology is a symptom of our culture, an observation difficult for many to accept. Heidegger (1977) has pointed out that technology is not only a means to an end, but that technology is a human activity. Truly, technology begins not with the tool, but with human imagination and creativity. Imagining, creativity, invention, and using tools and machines are human acts. Technology as a human activity is part of culture and reflects human preoccupations and concerns of the time.

Invention and innovation occur in a social context and milieu. As Rybczynski (1983) has pointed out in his excellent

book on technology, the optical lens presents a case in point. Lens were known to Muslims as early as the tenth century. Wide practical use (innovation) did not occur until four centuries later in Europe. A period that supported literature and reading brought the use of lens in reading glasses at the same time that printing came into use. Some three hundred years later a military innovation used lens in military telescopes for artillery spotting. This is a typical role of military innovation in technological development.

Bcok printing is a related example. Printing was not an object but a human process that extended human capabilities of writing and reading to provide insights and information. Human discovery was enhanced by printing, as will hopefully be the case with the computer in the technological environment of today.

Controlling material technology must not only consider the technical and mechanical factors, but the human factor as well. The example of the Lordstown experience of General Motors has already entered the folklore of industrial management. We hopefully learn that the challenge is more than generating a technology with a happy face, but to recognize and deal with the human component present in all technology, whether material or non-material.

Our attitudes toward machines often lead to a romantic nostalgia that may be dysfunctional for the future. Sailing ships are romantic, but lack the conveniences expected today. While the word processor is replacing the typewriter, many frown on the dot-matrix printer because the product doesn't appear "typewritten"! A similar case may exist regarding missiles and spacecraft in the Air Force because we are enamoured of airplanes. This could represent a serious crisis in the future management of aerospace technology.

Behavioral Guides for Managing Technology

Man's greatest discovery is the use of his mind in converting knowledge and science to invention and innovation in the form of tools. Managers are finding themselves as translators and multipliers of advancing technology in making decisions and solving problems to advance the welfare of mankind.

As a multiplier of technology, the manager uses organization and management to bring products of science and endeavor to the advantage of society. Managers today are vitally concerned with invention, innovation, and the diffusion of knowledge.

Stephens has described managers as conduits of technological change (1977). This managerial role involves the acquisition and processing of knowledge to extend man's intellectual capability by means of technology. This process is rife with uncertainty, complexity, and often lacks the rational order desirable in the scientific method.

The conclusion to be drawn is that many old rules of management and organization need modification for success in the Machine Age. Clearly, what always worked might not work in the future. The old folklore of "If it ain't broke, don't fix it" is most inadequate to prepare for management of technology in the future.

Further, the refinement of machines does continue with the excellent work done in laboratories and research centers. The Booz-Allen and Hamilton studies have attracted wide attention for their findings that point out we have good machine technology, but are not properly introducing it into organizations and operations. This can only be done by man. Machines can't introduce themselves into strategy and planning, let alone operations. Here again it is wise to remember the previous note that technology is a function of human activity.

Technological change must be accepted as a given. This applies to management style, training, job design, and communications. While accepting the benefits of change in terms of upgraded standards of living, the loss of security and comfort of nostalgic old practices provides a bittersweet outcome. It should be remembered that writing, printing, decimal numbers, and the discovery of money were also technological changes in their day.

Organizational implications of our technology probably will bring flatter organizations, broader jobs, upgraded management control systems, and emphasis on planning and resource use.

Behavioral science will be in the thick of the fray to use what we know about human behavior to advance technology as a human activity. As in the Booz-Allen studies, behavioral inputs will constitute an important segment of human technology in the more successful organizations. Our current best seller In Search of Excellence by Peters and Waterman reflects just that behavioral orientation.

Key Areas. Some vital areas of behavioral management of technology will follow. These areas provide opportunities for initiative.

Training will become even more important. The continual upgrading and maintenance of skills and capability will be aided by technology. However, the investment and planning for training will be one of the most fruitful areas of organizations. In lower technological organizations, training will mean survival. In the high technology organizations, training will provide the fine edge that augurs competitiveness.

Management style will be keyed to flexibility and standards. The standards will indicate levels of desired performance, while permitting flexibility in means or process. Flexibility will indicate sensitivity for change processes. This flexibility will accompany a search for improvement, denying the comfort of custom and habit.

Communication will be a key variable. The technology manager must be informed, and keep his relevant colleagues informed. Open and continuous channels will be maintained, while selectively establishing needs and priorities for information to preclude information overload.

A sense of "limits" of technology will give the technology manager an awareness of the returns to be generated from ventures or strategies. The technology limit curve interestingly approximates the ogive learning curve, and portrays the marginal productivity of additional technology inputs.

The culture of the organization will be recognized for its vital relationship to technological coping and competitiveness. The corporate culture consists of the values, priorities, history

and leadership climate. Successful organizations are seen to possess an enhancing culture which doesn't spring up overnight. The corporate culture will become a very important variable in providing the launching pad for invention, innovation, and diffusion.

Tolerance for risk, complexity, and uncertainty is a prime ingredient in managing technology. While related to the corporate culture component, this tolerance can be a personal behavioral variable as well as a factor in management style. This tolerance is a supporting element in promoting invention and innovation. The "high tech" industries are characterized by an atmosphere that encourages taking risks, bringing order out of complexity, and reducing uncertainty. This is the classic "uncertainty absorption" role of the systems manager. When mid-managers typically revert to risk-aversive tactics, it is most important for senior management to reaffirm the risk-taking ethos of the enterprise.

In summary, technology can be managed, but it requires a conscious awareness and direction of effort. Since technology is a human activity, managers will play a key role in stimulating and implementing the contributions and discoveries of scientists and researchers. We are talking about shaping human behavior to extend the capabilities of man through technology.

References

Booz-Allen, and Hamilton. Technology Management Survey Results. 1981.

Booz-Allen, and Hamilton. Technology Management Presentation Summary. 1982.

Drucker, P.F., Technology, Management, and Society. New York: Harper, 1970.

Heidegger, M.. The Question Concerning Technology and Other Essays. New York: Harper. 1977.

Rybczynski, W. Taming the Tiger. New York: Viking Press. 1983.

Schon, D.. Technology and Change. New York: Delacorte Press. 1967.

Stephens, J.C.. Managing Complexity. Mt. Airy,MD: Lomond Press. 1977.

AD-P003 350

Methodological Issues in the Study of Air Force Organizational Structures[1]

Captain Jeffrey S. Austin, Ph.D.
Major Lawrence O. Short, Ph.D.
USAF Leadership and Management Development Center
Edward J. Conlon, Ph.D.
University of Iowa
Richard L. Daft, Ph.D.
Texas A & M University

Abstract

Two theoretical models are presented for the study of organizational structure. These models are described and used in pilot work in two functional areas at eight Air Force installations. A method for collecting rich data which can be analyzed within each model is described. The theoretical and methodological frameworks allow researchers to look at vastly different structures and help "standardize" results.

Recent research literature indicates a great deal of information has emerged in the past ten years concerning various non-traditional approaches to designing organizations (see, for example, Daft, 1983). This interest has been sparked by an increased awareness of the environment and its affect on human resource management. Interest within the Air Force over the past few years has been highlighted by the numerous requests for research studies to investigate more efficient and effective ways of utilizing people.

The growing practice of using non-traditional organization structures is driven by the need to cope with rapidly changing technologies, unique customer requirements, and the need for multi-disciplinary teams to solve complex problems. Other key factors are financial and human resource constraints which impel managers to organize and utilize people in the most efficient way. Thus, changes in the political, social, and economic environments of organizations, coupled with the high rate of technological change, produce a turbulent environment. In this environment, an organization striving to maintain a dynamic relationship with forces in the environment finds that it cannot substantially reduce the risks and uncertainty under which it must operate. The consequence has been growing concern with developing structural forms in the organization that will adapt more easily (Davis, 1979).

---

[1] The authors wish to thank Major Richard Hayes for his contributions to this paper. Space limitations prevent a complete discussion of the LMDC organizational design pilot study referred to herein. A complete report on this study and its findings is available from LMDC/AN, Maxwell AFB, AL 36112. The opinions expressed in this paper are those of the authors and do not necessarily represent those of the U.S. Air Force.

Importance of Structure

When structure fits the needs of the organization, it is hardly noticed. However, when organization structure is incorrect, when it is out of alignment with organization needs, one or more of the following problems may appear (Child, 1977; Duncan, 1979).

1. The organization does not respond quickly or innovatively to environmental changes.
2. Too much conflict is evident.
3. Managerial decision making may be delayed or lacking in quality.
4. Employee motivation and morale may be depressed.
5. Resource utilization may be uneven.
6. The organization will not achieve performance goals.

These symptoms are indicative of the ways that structure impacts far reaching issues within organizations. Further, this impact seems apparent in either civilian or military organizations.

Air Force Perspective of Organization Structure

Functional grouping has been the traditional Air Force approach for many years. Air Force Regulation (AFR) 26-2, Organization Policy and Guidance (1982) emphasizes that the most effective functional groups are made up of functions that have a common goal. USAF principles, objectives, and policies indicate functional grouping is to be the predominant form of organization structure. However, in an organization as large and complex as the Air Force, the functional approach does not always apply, and variations or alternative approaches can be more effective. Regardless of the organization design, it is important to realize the basic objectives of the Air Force organization as outlined in AFR 26-2. These are to (1) maintain a structure that operates effectively with the least expenditure of resources; (2) standardize the organization structures as much as possible; (3) keep pace with technological advances, changing missions, and concepts of operation; (4) streamline the decision-making process; (5) ensure that the organization of improvements in one part of the Air Force are applied elsewhere, when applicable; and (6) to develop organizational nomenclature that has precise meaning throughout the Air Force.

HQ USAF/MPM (Headquarters, United States Air Force Manpower and Organization) has requested studies be conducted to investigate non-traditional methods of organizing people and skills. According to a letter from MPMZ, dated 4 Dec 81,

> "The Air Force does not have a capability to make an objective, a priori comparison of the advantages and disadvantages of alternate techniques of organizing and utilizing people to accomplish mission requirements."

Subsequently, the Leadership and Management Development Center (LMDC) initiated studies to address this problem. This paper reports on the issues of identifying models and developing methodologies for addressing this long term assignment by describing the design and conduct of the initial pilot study.

Description of the Models and Pilot Study

Two models were used for data collection purposes in the study. Each model is effective in understanding the complex framework within which structure is

determined. Future assessment will be conducted to select the most appropriate model for a given structure. On the surface, each model appears to assimilate important data, and it may be that a combination model may be the best solution.

Descriptive Models

The first model deals with a continuum for assessing alternative organizational structures. The structural allocation of responsibility, division of labor, and grouping of departments within organizations typically follow one of four structural forms. These structural forms are reflected in the design of the organization chart, and are typically identified as follows:

1. Functional structure. People and departments are grouped together by common functional activity.

2. Program structure. People and departments are grouped together by program, product, or geographical area.

3. Hybrid structure. Part of the organization has a functional structure and part has a product structure to gain advantages of each.

4. Matrix structure. Product and functional structures are implemented simultaneously and overlay one another. A complex form of structure used only for unique circumstances.

An underlying theme that governs the use of these structural alternatives is efficiency vs. effectiveness. Organization structure can either be slanted toward achieving internal efficiency or toward achieving effectiveness in response to external demands. Organizations that are structured to achieve internal efficiency typically exist in stable environments and have technologies that are routine and predictable. These organizations try to make the most efficient use of human resources. At the opposite extreme are organizations that must ignore internal efficiences in order to accommodate explicit demands for external effectiveness. These organizations must respond to changes in the external environment and work with the technologies that are non-routine and unpredictable. These organizations must be designed for innovation and coordination. Figure 1 illustrates a continuum anchored at each end by internal efficiency and external effectiveness. A specific structure should be adopted based upon advantages for the organization's specific needs.

The second model, adapted from Galbraith (1977), views structure as highly interdependent with a number of additional organizational characteristics. Like most modern perspectives on organizational design, this framework sees the structure utilized by an organization as a variable aspect of the organization's strategy. More specifically, the choices of how an organizational unit should be structured are part of the general strategic decision process and cannot be made independently of decisions regarding the organizations goals, the technology used to pursue those goals, the environments in which the organization will function and its people. The structure utilized by an organizational unit depends (or should depend), therefore, on other aspects of the corporate strategy. Galbraith identifies four such elements which are (1) goals and objectives, (2) environment, (3) human resources and (4) technology. The network of interrelationships is shown in Figure 2.

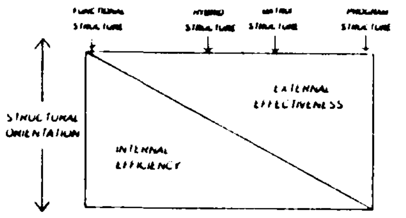


FIGURE 1 RELATIONSHIP OF ORGANIZATION STRUCTURE TO EFFICIENCY VS EFFECTIVENESS ORIENTATION

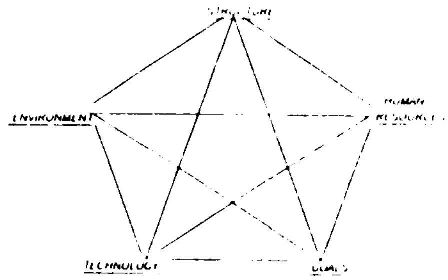


FIGURE 2 A SCHEMATIC OF THE GALBRAITH FRAMEWORK

Pilot Study

Data used in the initial pilot work were collected through a series of structured interviews across selected bases within several major Air Force commands. Four experienced interviewers were used during the data gathering phase. The interviews were standardized to the degree that two forms were developed to guide the interviews. Daily meetings were held to discuss problems, terminology, trends and common themes.

Two separate sets of data were collected. One set of data was focused within the Air Force aircraft maintenance community, and the Galbraith model served as our analysis paradigm. The interview solicited information concerning the nature of the job, the mission and associated goals, technology, interaction and communication patterns, climate and organization structure. In particular, structure questions focused on the differences in the organizational methods that currently exist in maintenance. Those interviewed were also requested to comment on particular strengths and weaknesses of the structures from their own viewpoint. Finally, effectiveness criteria were discussed.

The other major functional area that contributed to this effort was that associated with research and development organizations. The structures studied were varying forms of product/project management, most typically in some form of matrix structure. These data included information about the job and its associated complexity, mission and goals, amount of change, interaction and communication patterns, climate and structure. Particular emphasis was directed to interaction patterns in this fairly complex arrangement. Once again, we sought information regarding particular strengths and weaknesses and measures of effectiveness.

Interviews were scheduled for one hour, although the actual duration of the interviews varied. Generally, the lower in the organization, the less knowledge and experience the individuals had with the issue of structure. Nonetheless, their input was valuable in assessing job demands and communication interaction patterns.

Those interviewed included personnel from eight different sites. We interviewed a total of 74 personnel within the matrix structures and 106 people from varying maintenance structures. The people were selected by the research team by position to insure both representativeness and consistency across units. The selection was purposely stratified at a high management level to help quickly assertain more global issues and to assure widest amounts of varying experience in organizations.

Discussion and Conclusions

We have considered several issues in this paper regarding organization structure. Included are theoretical models for structural variations and alternatives and a methodology which can be used to study structural issues. The use of this methodology and theoretical foundations would seem to provide advantages when studying organizational design and structure.

First, use of the models and methodology form a process which provides an opportunity to look at organizational structure on a general scope. It allows the researcher to look at vastly different types of structures and to make comparisons as a matter of course. The theoretical foundations provide a convenient way not only to "standardize" results but to consider the application of one type of structure to another organization and anticipate positives and negatives ahead of implementation. In addition, the process is one which may be widely applied within the Air Force. While space was not available to present results of the current LMDC organizational design study, the application of the process provided a framework within which to make future organizational structure decisions. This is an opportunity which has apparently been lacking in the Air Force and one which may be fairly easily done by a wide range of Air Force people.

In the final analysis, the ease of application may be the biggest advantage of the process. It provides a "roadmap" of theory and practice which others may follow. All too often, we seem to create structures which people must work around to get the job done. If we are to "do more with less," structure seems a crucial place to start in helping people be more efficient and effective. The process outlined here is meant to help Air Force people do precisely that - become more efficient and effective.

References

Child, J. Organization: A guide to problems and practice. New York: Harper and Row, 1977.

Daft, R. L. Organization theory and design. St. Paul, MN: West, 1983.

Davis, L. E. Optimizing organization plant design: A complementary structure for technical and social systems. Organizational Dynamics, 8(2), 1979, 3-15.

Duncan, R. What is the right organization structure? Organizational Dynamics, 1979, 7(3), 59-80.

Galbraith, J. R. Organization design. Reading, MA: Addison Wesley, 1977.

Leadership, Managership, and Computers in Today's Air Force

Lieutenant Colonel William E. McCarron and Captain David R. Harvey

United States Air Force Academy

Abstract

Managership is quantifiable and measureable while leadership is qualitative and personal. Though the two concepts are not opposed to each other, there are differences. This paper discusses specific Air Force situations to show that leadership extends beyond managership, particularly in computer applications areas. For a mere manager, computers make decisions; for a leader a computer is only one tool in decision-making. Exclusive managership leads to narrow thinking and predictable tactics. Leadership includes people, weapons, and a consequent dynamism which yields innovative thinking and tactics.

McCarron: A visitor to our Air War College seminar recounted a personal experience that well illustrates the difference between leadership and managership in the Air Force. When the first operational F-16 landed at Hill Air Force Base, Utah, half the base came out to welcome the pilot and the airplane. A "manager," viewing the situation, observed, "Don't these people have something better to do with their time?" A leader, viewing the same situation, said, "Where's the other half of the base?"

Although leadership and managership are not necessarily opposed to one another, there are distinct differences. Managership is appealing because it is quantifiable, measurable, tangible. One can just see a military manager calculating the hours lost as civilian and military workers at Hill gathered around that electric jet. Leadership is more difficult to detect: it is qualitative, earned, personal. You can't add up leadership hours and include them on a slide for the wing mission briefing at Hill AFB, but you know if leadership is there, or if it is not.

Any military commander has authority because it comes with the job; he or she has it by virtue of the office whether that office be the base commander, a squadron executive officer, or a crew chief. We presume that a commander is a good manager because that is how a commander will be judged on a daily basis. Management of resources is the military hierarchy's ultimate measuring stick for how well a unit is doing, and those resources include the traditional three "Ms": material, money, and men. Every inspection, or combat sortie, or military exercise reduces itself to how well did the unit use its equipment, in terms of dollars expended, to achieve an objective. The competent commander and battlefield manager will be able to click off responses about men, money, and material quicker than a supply officer can quote from an inventory sheet. That's as it should be, particularly in the age of computers, distributed information systems, and word processing.

Increasingly in the modern world, computer technology drives military managership. For example, high-speed integrated circuits and microprocessors currently enable the Air Force to construct an air combat maneuvering instrumentation system. Fighter pilots are literally able to see their missions recreated in large video-game format thanks to Heads-

Up Display cameras and advanced instrumentation. The shapes moving around the screen are not Space Invaders or Pac Man, but recreations on film in three dimensions. What a management tool! The guesswork is now taken out of basic fighter maneuvering evaluations. Gun tracks, missile shoot cues, movement indexes, and G-loading are there on the debriefing screen for the fighter pilots to review.

There is, however, an inherent danger in literally such "micro management." The commander who is management-oriented can concentrate too exclusively on the quantitative and can forget the qualitative. Of the three resources mentioned above, the military traditionally considers the people resource the most important. In quantitative terms, the Air Force can manage money and material, but a strictly quantitative managing of people is another matter. It is as if members of a military unit were so many pieces on a war gaming board. The whole concept is too impersonal. We lead people, or we follow them; we don't manage them. To manage people is to conceive of them as pawns on a checkerboard; to lead them is to become part of the game—to join them on the war gaming board—to become part of the action.

Harvey: In a gesture that is humorously symptomatic of concentration of the finite and the measurable, Time magazine recently named "the computer" as their 1982 "Man of the Year." Desk-top computers are now nearly as common to Air Force homes and offices as file cabinets and telephones. Information exchange and retrieval are the orders of the day. For example, each F-16 aircraft maintenance unit in the Air Force has a dedicated computer which shows the status of each F-16 line replaceable unit (LRU) in the Air Force inventory. Thus, Air Force managers can inquire worldwide about needed LRU's with the punch of a few keys. The computer is a distinct benefit as a management tool, but it is not necessarily a leadership tool. The computer can yield a matrix for decision theory, but it must not make the decision. Leaders do that. Computers have logic gates, intergrated circuits, huge memories on microchips a fraction the size of a human fingernail. Leaders can never manipulate information with the rapidity of a computer, but only leaders can achieve insight into human problems and make connections that convey loyalty, honesty, and integrity. Computers can digest discrete bits of information; only a leader has the qualitative insight to know when and how that information serves a mission, a unit, and the people who work there.

It may be a bit unfair to set the computer up as a strawman for the issue of leadership versus managership. Taken by itself, the computer is an indifferent tool. As Aristotle reminds us, tools are only material causes; thus, computers take on meaning in relation to how a man or woman uses them. Computers help in managing money and material, not men. But even then, there are dangers—like when the automated system becomes more important than the problem you're trying to solve. My first assignment in the Air Force was at a large intelligence wing. The wing had been using a intelligence data handling system since the early seventies and was then in the process of replacing it. The system had started out as a simple system designed to meet the needs of imagery interpretation and collection. But as the decade went on, more functions were added to the system, till it became so complex that it wasn't meeting the needs of even the imagery interpreters. But like most large systems, people learned to live with its inadequacies. The chance to replace the system in the early eighties, however, provided a great opportunity to correct these shortcomings.

Accordingly, the command hired consultants to come into the building and study our situation and make recommendations on what type of new system we might need. This

was a good idea, it seemed, but I started to get nervous when the consultants never came around to our work area. After all, we were the ones using the system. I found out, years later in a course I took on implementing information systems, they were using the "innovative apporach" to designing a system. That is, they refused to look at the older system so their ideas and solutions to our requirements would be fresh, unhampered by the old, outdated approaches we'd developed. But how can you design a system to solve a problem when you don't know what the problem is? In our command the situation was made worse because the consultants didn't have the proper security clearance to gain access to the materials we were working on. Thus, they couldn't even talk about our problems, let alone solve them.

The situation came to a head months later while I was on the Source Selection Evaluation Board at the Air Forces Computer Acquisition Center at Hanscom AFB. I spent a whole month reading the proposals that came in bidding for the contract. The proposals for the most part were exactly what the Air Force, at the consultants' request, had asked for. All the companies proposed large, state-of-the-art mainframes with supersophisticated graphics. But as I sat there reading, I took a pad of paper and started writing down questions. Later, I'd approach one of the computer technicians and fire away. "Why do our requirements ask for all these graphics?" "Graphics, son, are the most exciting area of computer design these days; they're state of the art." "Oh, but we don't use graphics." "Well, now you're gonna have to, I guess. Excuse me. . ." The next big question was, "Will photo interpreters be able to access the master list of installations?" "No, I'm afraid not."

Such managership situations are not uncommon and they're costing the Air Force millions. And the problem is not unique to SAC, nor any other command in the Air Force. Until we put computers in their place, as mere tools, until we stop worshipping at the shrine of technical expertise, and force the technicians to pay attention to the users, until we stop oohing and ahhing at color graphics when what we need is a simple filing system, and until we learn to treat the machines in the same way our secretaries do their typewriters, computers run the risk of costing us even more, and still not doing the job.

One version of a computer familiar to many Air Force people today is a word processor. Take the situation in our department's front office for example. We have one head secretary and four educational technicians (typists). Each has a typewriter, but we also have a word processor. The AD types call it a unique, stand-alone, dedicated word processor. It works great! All of our clerical personnel have learned to use it. The Academy, however, has been forging its way into the "information age" and wants to put a word processor in every department. Great idea. I couldn't agree more.

But then I found out the resources for the new word processors had been put on hold until AD could find a suitable distributed data base system. A distributed data base is different because the department won't have control over it. I was incredulous. "Why can't the departments have what they know will satisfy their needs?" An information systems manager would answer, of course, "The user doesn't know what he needs." But I think behind his snobbery is a real fear: he's afraid that the user does know what he needs, and what he doesn't need is the information system manager. And that's the point. Too often in the Air Force, or any other organization, the desire to automate takes on a life of its own. It grows in cost and complexity till the computer types are secured in a job, and the users are either confused or have forgotten what their original needs are, or

they are unsatisfied. And as for getting information to the manager or the leader, well, I don't think that's the issue any more.

McCarron: The problem of worshipping at the shrine of computer technology and ignoring the needs of the congregation is not, by any means, the exclusive bailiwick of the Air Force. Industry, government agencies, universities are battling the same problem. The syndrome of the "the system is the answer" just isn't true. It ignores the people who must see if the system meets their needs. We're conscious of this fact; the Europeans even have a fancy name for it: ergonomics, or human factors engineering. The Air Force, industry, private corporations are interested in how people interface with the system. But, ergonomics goes both ways; it is also how the system interfaces with the people. And lots of people are struggling with this second half of the ergonomics equation!

A manager says, "A computer solves my problems." A leader says, "A computer helps me solve some of my problems." A manager is left sucking silicon bytes. A leader knows when to wean himself or herself from the machine. A manager depends on high technology. A leader depends on what John Naisbitt (1982) calls "high tech/high touch": "When we fall into the trap of believing, or more accurately, hoping that technology will solve all our problems, we are actually abdicating the high touch of personal responsibility."

If there is any one thing that distinguishes a leader from a manager—whether the battlefield be actual or a problem with computers--it is a willingness to mix and maneuver with the troops, to reach a deep understanding of people problems. One doesn't have to wear ivory-handled pistols, nor wave floppy discs. The less ostentatious, the better. A friend of mine recently assumed command of an aircraft generation squadron, and the pressures were on him from day one to implement a better management system, to get rid of the "hangar queens" that occupied space along the flightline, and to improve the sortie generation rate. He did all these things because he was a solid manager; he did more, though, because he was a leader. He takes each Friday afternoon off and wanders through his aircraft maintenance units, pausing to talk to a young airman who has just returned from a two-week deployment with the F-5 Aggressors or stopping to discuss a problem some of his NCOs are having keeping the tech orders up to date. It means he has to work a Saturday morning and it means night work a few days a week but, as he puts it, "It gives me a chance to visit with the night crews before I head home. They're the ones getting the planes ready for Red Flag the next day."

Whatever a leader's association with the troops, it cannot be mere tokenism. It must be substantial. The people who follow you must do so not to elevate you, but because they know you care about them and the mission at one and the same time. Leadership is genuine concern, not a contest to see who the most popular colonel on base is. Popularity contests are for fashion magazines and Hollywood where how one looks counts more than what one is as a person. Leadership is fatigues and crawling under an airplane with a "slick sleeve" airman and letting him or her tell you what pneudraulics repair is all about. It's a phone call home to a wife or husband asking to have dinner kept warm in the oven because you have some problems that need attention, not just some papers to shuffle or some bar talk to exchange. Leadership is knowing how to tackle deep squadron or wing problems, not just manage their surface solutions.

Leadership in the modern Air Force is rarely sensational and on the electronic battlefield, it's unlikely that a Patton or MacArthur will emerge quite as readily as in the

past. Whatever leaders emerge will be people who know their people because a leader is first a leader and then a manager. Still, what the military demands of us most of the time is solid management: so many flying hours logged, so many flight control computers on line, so many successful launches of AIM-7's and AIM-9's. But, today's Air Force isn't just a business, measured in terms of a Dow-Jones Industrial average. The rise or fall of the fortunes of war will depend on how well each of us knows our people. A computer may be Time Magazine's "man" of the year, but Air Force men and women are leaders and followers each day of the year.

Harvey: I've been taking courses on managing information systems for the last year and can't count the times I've been reminded to beware of the "non-technical, computer illiterate user." But the term "user" sounds like it belongs in a courts-martial proceeding, and what the teachers and texts and students are forgetting is that it's these "users" who have the real job that the system is supposed to make eaiser. The "user," mainly a high school educated airman or a young officer who would love to do a professional job, becomes frustrated using a computer system that was designed to be impressive, but not to solve his problems. It doesn't take a lot of experience and education to see how much money was put into a machine to have pretty graphics and how useless it is in his own job. And when it becomes obvious to the airman he's there to serve the machine rather than the other way round, he loses interest and the quality of his work suffers. I saw it happen hundreds of times during my first assignment. Bright, hardworking young men and women came into the outfit, spent a few months working on the computer to produce intelligence reports, and finally lost interest because the computer was an obstacle rather than a tool, because it looked nice in briefings, but the briefings never covered what the real job was. And they developed an instinctive distrust of the system—not just the computer system, but the whole system, the Air Force -and knew that the new computer wasn't really going to be any different. And, you know, they were right. The final tragedy, of course, is that the Air Force loses these airmen and officers. The moral of the story is simple and information system leaders need to take heed: If you treat your men like USERs, they'll act like USERs; if they're treated like people with a job to do and the computer system is made to solve their problems—not the system designers'—the job will get done.

McCarron: It is precisely a leader's concern for people each day of the year which underlies the comments of General Gabriel (1983): "In reality, our "magic weapon" is our people—people who are well-trained, who seek and practice innovative tactics, and who apply our Country's technological advantages. . . .We have a superb Air Force today built upon a sound qualitative foundation and will make it even better tomorrow with innovative, disciplined thinking, and strong leadership". For a leader, people are "magic weapons"; they count more than the high tech weapon systems—including computers—the managers love to inventory. For a leader, innovative tactics win in combat, not just approved tactics which are safe but predictable. Leaders transcend managers on the vital issues. They advocate weapon systems and tactics that are people-oriented and geared to win. Managers count; leaders have people they can count on. Managers measure; leaders make sure their people measure up.

References

Gabriel, Gen. C.A. Training and flexibility: our keys to the future. TIG Brief, 1983, 1, 2.

Naisbitt, J. Megatrends. New York: Warner Books, 1982.

Leadership on the Future Battlefield: Systems-Wide Perspectives

J. G. Hunt and John D. Blair
Texas Tech University

Abstract

The battlefield of the future is described and seen to be chotic and stress invoking to a degree not encountered heretofor. A heuristic systems-wide model is developed to deal with arms leadership and organizational requirements imposed by the battlefield.

A number of works included in an upcoming book dealing with the above topics are briefly discussed within the context of this systems-wide model. These papers help form the basis for a much needed research roadmap to assist the Army in the extensive change efforts required.

Leadership on the Future Battlefield

The battlefield of the future is predicted to be one of enormous destruction resulting in great confusion and high levels of fear among all involved. The Army has developed the Air-Land Battle (ALB) 2000 concept which attempts to look ahead to the Year 2000 and portray what the battlefield would look like in a major confrontation in Europe between NATO and Warsaw Pact forces.

Conventional weapons have become far more lethal than before, and the possibility of chemical and tactical nuclear warfare is very real. Electronic warfare may make communication between units in the field and their commanders at headquarters impossible. Even attempts to communicate may result in destruction from weapon systems that lock on to radio signals. Soldiers will have to fight continuously with little or no rest because of night fighting capabilities. Rear areas, which have normally been relatively secure, are likely to be effectively attacked, and battle lines will be very unclear with both sides operating behind each other's lines with substantial forces.

What are the implications of this type of battlefield for soldiers and their commanders who must lead them and manage their activities and resources, and for the Army organization which must be designed and structured to function under these turbulent circumstances? These are among the issues addressed by several authors whose work we will summarize here. Their work will appear as chapters in the book we are currently editing on leadership on the future battlefield (Hunt & Blair, in press).

There is a cross-cutting integration of these works. Their theoretical and research efforts are generally focused within a heuristic systems model, which is outlined below. It includes organizational and environmental factors, leadership and managerial behavior factors, unit and individual subordinate factors and a range of effectiveness outcomes. In the following sections, we describe the overall conceptual perspective of the book, and report on these attempts to examine specific issues from a system-wide perspective.

Leadership on the Future Battlefield: A Conceptual Frame of Reference

Neither past traditional academic nor military definitions of concepts like

leadership or management should constrain or limit the perspective that researchers bring to their analyses. Leadership on the future battlefield is an extremely complex topic and requires one to go beyond the concerns of most traditional leadership studies. Here, leadership also includes managerial behavior, and both are examined in the context of the organization and its environment. Thus, there is room for broader concerns from the subfields of organizational behavior and organization theory as well as human resource management. In addition, leadership and management can be seen as different and, at times antithetical. However, both types of behaviors must be included since troops must not only be led in battle, but the organizational systems within which they fight must also be managed.

We have taken a very broad perspective on the range of variables and issues that should be examined under the general topic of future battlefield leadership. Such variables include macro and micro contingency factors affecting leadership and managerial behavior and the consequences of those behaviors for individual and unit outcomes. The contingency factors also directly affect the same individual and unit outcomes.

The model of these factors is shown in Figure 1. It indicates the major categories we consider important but is only suggestive of all the individual variables that make up each general factor. Here leadership on the future battlefield is seen as a highly complex and interrelated set of dynamic processes involving an open system with extensive feedback processes. By attempts to convey the extensive range of factors which will likely influence leadership on the future battlefield, the model thus points to the breadth of specific topics necessary to examine systematically the general topic. Of course, the ultimate challenge is to develop a very parsimonius model of the key variables.

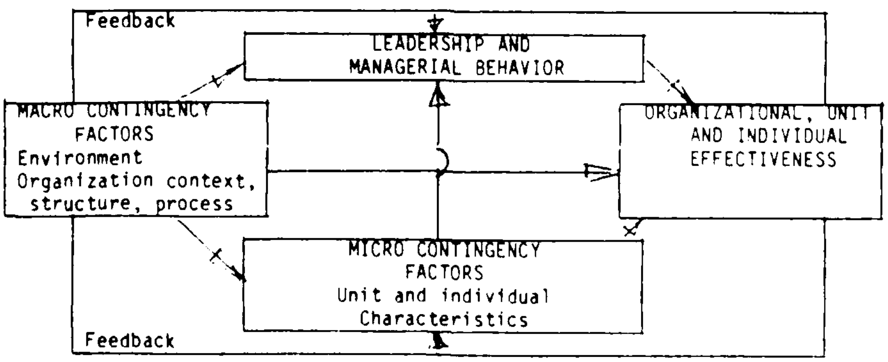


Figure 1. A Heuristic Model of Leadership on the Future Battlefield

Current Research from a Systems-Wide Perspective

The heuristic model argues strongly for a systems-wide perspective in looking at future battlefield theory and research. Current Army policy and doctrine is also consistent with this thrust. This in itself is a significant move forward for much of the current leadership research lacks a consideration of the context within which it is embedded (Haythorn, in press; Osborn, in press; Tosi, in press). A key problem with much of the Army's earlier [illegible] was that it tended to be restricted to narrow segments and lower [illegible] levels. Recognizing this, the Army is now beginning to express

doctrine for the future in terms of systems concerns. The development of leadership policy and doctrine is elaborated by Grunstad (in press). Thus, the managerial/leadership behavior breadth and depth issues addressed by Hunt (in press) and by Van Fleet and Yukl (in press) are treated, and the importance of synthesis as opposed to analysis is recognized by both Osborn (in press) and Clement (in press).

The systems approach to leadership concentrates heavily on senior level leaders and argues that to meet the Army's difference in kind (as opposed to difference in degree) leadership requirements for the future, senior Army leaders will need to be involved with the following organizational activities in order to achieve the multiple outputs of productivity, adaptability and stability.

First, they are responsible for setting Army vision--creating the future to be dealt with. This moves then into symbolic, charismatic, and ideological kinds of concerns as articulated by Hunt (in press), by Segal (in press), and by Van Fleet and Yukl (in press), and calls into question existing organizational paradigms as discussed by Whitehead & Blair (in press).

Second, they need to design interdependencies among appropriate organizational subsystems. These interdependencies are important because of anticipated future battlefield conditions. Involved here is developing control mechanisms so that individuals and subsystems can provide their own control on the battlefield. Thus, understanding the intent of the commander and being able to function autonomously are important in a turbulent battlefield environment with unrealiable command, control, communications, and intelligence (C3I) capabilities.

Third, senior leaders will need to create or recreate the organizational climate or culture. Five dimensions are envisioned here: supportiveness; trust, confidence and credibility; participative decision-making; openness and goal clarity among all concerned.

Last, Army leaders will need to establish information systems. This involves making sure that information processing functions are being adequately performed, that people are the key in an information processing system and that adequate feedback mechanisms exist throughout the organization.

The above four subtasks form the basis for planning and action and emphasize synthesis and integration rather than analysis and differentiation skills. The former are argued to be the dominant thought processes at the senior management level. Systems leadership is seen as providing support to the central concept of skills varying by organization level.

Concern with Strategy

Again, with a senior level emphasis and systems leadership is a concern with strategy. This also focuses on the macro and managerial behavior/leadership factors in the model. A particularly interesting approach involves analogs from corporate planning presented by Barton and Baliga (in press). The thesis is that even though leadership requirements differ by level, conceptual frameworks for decision making should be as similar as possible at all levels to help ensure the "intent of the commander" under the chaotic future battle-

field. The frameworks are designed to help deal with extremely high complexity and uncertainty. The notion of "screens" is introduced through which organizations can be viewed by people at different organizational levels in order to provide a common strategic frame of reference. The authors use corporate analogs and tie these to Army requirements to illustrate their argument.

Concern with Organization Design

Besides an emphasis on strategy, senior leaders will need to be concerned with organizational designs to enhance the concepts of systems leadership. Thus, the traditional mechanistic, bureaucratic organizational structure which has served modern armies successfully will no longer be appropriate. Whitehead and Blair (in press) argue that an organizational paradigm shift will be called for. This shift will mean thinking "paradigmatically" by those in the Army with mixed mechanistic and organic structures needed to meet the highly uncertain and complex future requirements.

Paradigmatic thinking is important because, without it, individuals are likely to be too wed to traditional mechanistic structures and modes of thinking which will no longer be sufficient by themselves. Thus, like Clement, (in press), these authors emphasize a way of thinking that allowsfor the kind of future organizational flexibility needed. Command, decisional, authority, control, reward and informational structures throughout the organization--as well as perspective, culture and behavior of the organizational members--will have to be managed.

In a companion piece, Blair and Whitehead (in press) conceptualize a systems-wide change model to be used to help move senior leaders and subordinates at all levels to the kind of paradigmatic thinking outlined above. In particular, the authors focus on developing ongoing systems-wide capabilities to innovate and adapt, i.e., to learn how to learn.

Osborn (in press) also discusses the kind of organization needed to deal with the complex, uncertain, future battlefield. He uses Weick's notion of tight and loose coupling along with organizational diversity notions called for by varying battlefield conditions. He links these with leadership requirements by providing examples of two kinds of coupling and diversity: tight coupling/little diversity and loose coupling/ extensive diversity. Leadership is conceptualized in terms of instrumental dimensions versus collective leadership emphasis involving value consensus. Finally, he stresses emphasizes the differences between the kind of collective value consensus leadership called for at senior levels and other quite different kinds of leadership needed at lower levels. For Osborn, as others, leadership is not immutable (varying only by degree); there are differences in kind, and not just degree, between senior and lower levels.

Tosi (in press) again contributes to this organizational perspective by proposing a model that examines the relative impact of formalization, technology, socialization, selection, reward systems, work relationships and leadership on predictability of behavior at different levels in organic and mechanistic organizations. He reminds us again, that with systems thinking many factors besides leadership are likely to be important. Furthermore, the relative importance will differ depending on level and type of organization. He then

extends the discussion into future battlefield scenarios.

Isenberg (in press) helps round out a treatment of systems-wide senior level leaders through his nontraditional treatment of what these individuals do and the skills they are likely to need. Based on systematic observations of top level civilian executives he concludes there is: (1) an absence of clearly discernable decision-making; (2) a heavy reliance on intuition; (3) a heavy reliance on process; (4) an emphasis on managing problem networks; (5) an emphasis on tolerating ambiguity; and (6) a premium on perceiving and understanding novelty. The contrast between the senior leader here and that revealed in Haythorn *et. al's* (in press) review of existing senior leadership research work is quite dramatic. Isenberg includes a very innovative discussion of battlefield implications of these observations and briefly treats selection, training, and development and discusses the importance of "managing decision contexts."

Concluding Concerns

As a wrap up, Segal (in press), among other things, reminds us that the future battlefield scenario outlined here may be only one mission in which the future army is likely to engage. The peacekeeping role appears to be an increasingly common one (e.g., Lebanon and the Sinai). He points to this and indicates how it is quite different from the ALB 2000 concept. These differing missions also remind us that, not only does there need to be a concern with organization and leadership on the hi-tech battlefield, but the Army must also function effectively in peacetime and quasi-peacetime as well. Segal and Blair and Whitehead also point to the uncertainty of what ALB 2000 will look like and to the fact that further changes can be anticipated.

Finally, Segal emphasizes the importance of environmental factors as he discusses implications of societal forces on selection and development of individuals for the future Army.

Discussion

There is perhaps no more important issue for the military today than how to prepare for the future battlefield. ALB 2000 tells us that there is less than twenty years to deal with far reaching changes. Indeed, the time is so short, and the changes are so important, that the Army has begun the first steps toward massive change largely without using research (Grunstad, in press). Such research in its current state, even with its limitations, has the potential to make substantial contributions to the Army if it is appropriately used. Thus, we have summarized a number of works by a range of scholars from several disciplines and subspecialties. However, all of their works point to the importance of conceptualizing leadership as a systems-level phenomenon that both affects and is impacted by macro-level factors and to the overall importance of adaptability and learning how to learn. These issues and implications, can serve as the departure point for a research roadmap which can focus on systems-wide issues of importance in moving toward this future Army. There are also a number of micro concerns discussed by the authors along with these systems-wide ones. Space has precluded their treatment here.

REFERENCES (Furnished upon Request)

A SYSTEMATIC COMPARATIVE ANALYSIS OF CONDITIONS ANALOGOUS TO LONG-DURATION SPACE MISSIONS

Jack Stuster, Ph.D.
Senior Scientist
Anacapa Sciences, Inc.

Abstract

This paper presents interim results of a NASA/Ames Research Center study to systematically analyze conditions of isolation and confinement analogous to a proposed space station. A methodology has been developed to evaluate the relevance of candidate analogues in terms of 14 dimensions or variables. Candidate analogues include research vessels, military outposts, offshore oil platforms, long-distance yacht voyages, and remote scientific stations, to name a few. Using the comparative method and focusing on critical incidents, we explore the behavioral, psychological, and social issues affecting human adaptation and productivity in isolation and confinement. The objective of the research is to develop specific design guidelines to enhance human productivity during long-duration space missions.

INTRODUCTION

Many attempts have been made to draw behavioral inferences from situations simulating long-duration space flight. However, with the exception of Sell's (1973) attempt to develop a taxonomy of confinement and isolation, little attention has been given to the relative **appropriateness** of the analogues, or to the likely utility of the inferences. Although the comparative method has been recognized by several investigators as a potentially valuable source of data, there has been little attempt to discriminate between the relative values of the many alternative analogues. There are several problems associated with **a priori** judgments in this area. For instance, a submarine making a 90-day submerged voyage is very similar to the proposed space station in length of tour and perhaps in the hostility of the outside environment, but fundamental behavioral differences may arise as a consequence of substantially different crew sizes or other dissimilarities.

Several features distinguish our approach from other efforts to study behavior in the naturally occuring laboratories of human experience. Most notably, our approach avoids **a priori** judgments regarding the relative merit of the many conditions of isolation and confinement which may be compared to the proposed space station. Others have assumed that underwater habitats, submarines, or Antarctic research stations, to name a few, provide good examples from which to extrapolate concerning human behavior aboard a space station. We make no such assumptions. Instead we have developed a methodology to evaluate the "relative degree of relatedness" of several candidates, or alternative analogues.

In order to evaluate alternative analogues, it was necessary to first prepare a list of dimensions, or metrics, to be used to define space station conditions. In developing these definitions, we were interested in establishing the parameters or assumptions concerning expected onboard conditions; these are required to allow the comparative evaluation of alternative analogues. To satisfy this objective, we compiled the following list of dimensions.

- Size of group
- Type of tasks
- Perceived risk
- Duration of tour
- Physical isolation
- Personal motivation
- Amount of free time
- Composition of group
- Psychological isolation
- Preparedness for mission
- Physical quality of habitat
- Form of social organization
- Hostility of outside environment
- Quality of life support conditions

Certain of these dimensions are, in fact, variables rather than givens. For instance, the quality of food is considered within the dimension, "Quality of Life Support Conditions." Clearly, the quality of food that may ultimately be served aboard the proposed space station--along with many similar decisions--has certainly not yet been determined. Also, in apparent contradiction, studies such as the current effort may be used to affect determinations regarding issues such as food. It is important, however, at an early stage of this project to specify conditions as completely as possible for purposes of comparison with alternative analogues. Even the dimensions which are clearly variable require some degree of specification--if only in general terms--in order to allow a systematic comparative effort.

A research instrument was developed which included a restatement of our list of 14 dimensions and definitions. This description of a NASA space station was followed by descriptions of several alternative analogues. The summaries of the analogues used the same dimensions (e.g., size of group, type of tasks, etc.) as the space station summary to describe conditions.

The Evaluation Methodology. The space station definitions were used as target conditions against which alternative analogues could be compared. This evaluation effort involved a dimension-by-dimension comparison and the use of a seven-point scale. Data collection sheets were provided for those participating in the study to record their evaluations; each data collection sheet concerned a different descriptive dimension. Together the data collection sheets formed the matrix reproduced in Figure 1.

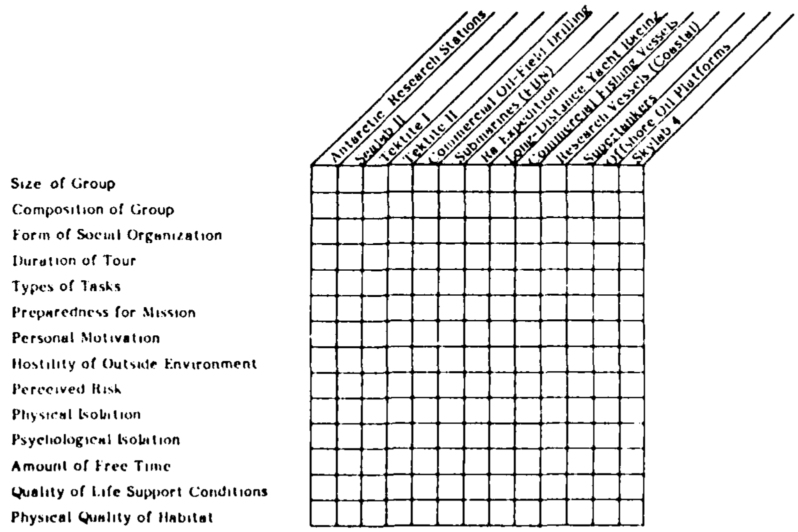


Figure 1. Matrix formed by combining data collection sheets (Appendix A).

An example of the seven-point scale is provided as Figure 2.

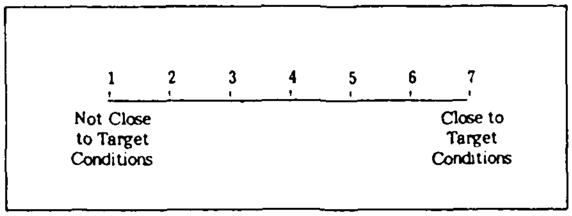


Figure 2. Seven-point scale used when evaluating "relatedness" of alternative analogues.

The procedure followed by evaluators was rather simple. For example, our space station summary assumed a group size of eight personnel. Those participating in the evaluation effort were required to select numbers on the seven-point scale to represent how closely they believed each alternative analogue corresponded to the assumed space station condition. An analogue with a group size of nine might receive a relatively high score on that dimension (corresponding closely to the crew of eight assumed for the space station). A group size of 90 would probably receive a relatively low score. Certain dimensions were more problematical than others--for instance, when comparing the physical quality of an analogue's habitat to that estimated for the proposed space station. Judgments such as these were made and recorded on the data collection sheets for each alternative analogue in terms of each of the 14 dimensions.

We recognize the difficulties inherent in quantifying subjective phenomena. We believe, however, that by providing systematic descriptions and then combining the judgments of many behavioral scientists and design engineers, we apply a more scientific method and, consequently, transcend the customary anecdotal approach to this subject. A systematic comparative effort is important because the results of the evaluation are being used to guide our efforts and determine our focus in a subsequent phase of the project.

Statistical analyses of the results of the evaluation have been conducted to identify those alternative analogues with greatest overall fidelity to expected space station conditions. The evaluation has also indicated which analogues correspond closely to expected conditions in terms of specific dimensions. Further research activity will be concentrated on those analogues indicated to be most promising by the analysis. By applying quantitative measures to comparisons of somewhat subjective conditions, we derive empirical rather than solely intuitive measures of relative analogy.

What follows is the summary of space station conditions included in the research instrument. Results of the evaluation and design guidelines were not available at the time this paper was prepared. Preliminary results were discussed, however, during the presentation of the paper.

NASA SPACE STATION (ASSUMPTIONS)

Introduction. It is likely that the construction of the space station will be an evolutionary process. In the earliest phases it may be a single cylinder attached to an enormous wing-like array of solar cells. Gradually, additional modules will be added until the space station will appear, from the flight deck of an approaching shuttle, as a grand, high-tech tinkertoy. It will not conform to our popular conceptions of what a space station should be. It will lack the lyric quality of orbiting stations depicted in novels and films; there will be no gleaming giant wheels rotating to the pleasant strains of a Strauss waltz. Rather, by the closing years of this century, the U.S. space station

will be a busy factory in the sky. It will be, first and foremost, a place of work. Onboard operations will likely involve facility and satellite maintenance, astronomy, basic science, and the commercial production of precious commodities.

Size of Group. Since the construction of a NASA space station would be an evolutionary process, we have selected the range of 6 to 12 persons for our definition. We assume a resident crew of 8 within two years of station deployment.

Composition of Group. It is expected that composition of station crews under routine conditions will be somewhat mixed in terms of sex, age, ethnicity, education, and work history.

Form of Social Organization. It is assumed that the form of social organization that has evolved within NASA for STS missions will be applied to the organization of work aboard a NASA space station. That is, a quasi-military structure with a commander, mission specialists, and payload specialists.

Duration of Tour. We anticipate tours of 60 to 90 days under operational conditions. Schedules of personnel rotation cannot be specified at this time.

Type of Tasks. Although specific information regarding the tasks involved in zero-gravity electrophoresis and materials processing are clouded by proprietary issues, we may safely assume that most onboard tasks performed by station crew will be of a vigilant and hand manipulative nature. Repair and replacement of components may be a frequent function. Extra vehicular activity (EVA) to service unmanned platforms and satellites, which is quite strenuous, will also be required.

Preparedness for Mission. It is expected that a great degree of preparation for space missions will continue to play a substantial, yet diminishing role in the future. Space station crews are likely to be at the extreme on this dimension compared to all other analogous conditions.

Personal Motivation. It is anticipated that there will be many more volunteers for positions aboard a space station than there will be positions available. It is also assumed that government pay scales are not primary motivating factors for application. For these and other reasons, it is assumed that the personal motivation of crew personnel will be other than financial.

Hostility of Outside Environment. Without mechanical means, human life cannot be supported in the environment outside the space station.

Perceived Risk. Exposure to risk will be substantial. In addition to the risk of system failures, we must consider the potential for micrometeorite collision, solar flare danger, and critical human error. The possibility exists that personnel will be required to spend a maximum of 21 days in an onboard "safe haven" awaiting rescue from a catastrophic incident.

Physical Isolation. Since the proposed NASA space station will be occupying a low earth orbit, the physical isolation from the outside world will be complete.

Psychological Isolation. It is frequently impossible to separate physical from psychological isolation. In the case of a space station, we assume that the capacity to communicate with ground control personnel and even with family members will be allowed in order to reduce the crews' feeling of psychological isolation. We anticipate that periodic, scheduled calls home will be a part of routine station operation.

Amount of Free Time. Although this dimension is of a clearly variable nature, we feel that it is necessary to estimate the amount of free time available for purposes of comparison with analogous conditions. For instance, there appears to be an abundance of free time during Antarctic winters, but very little is expected onboard a space station. Based upon the need for maintaining high levels of productivity to justify costs and on the experiences of previous space missions, we assume approximately 2½ hours per day will be available for recreational pursuits.

Quality of Life Support Systems. It is assumed that the atmospheric pressure would be maintained at 14.7 psi, the same as standard sea-level conditions; the atmosphere would likely consist of 79% nitrogen and 21% oxygen, again similar to earth conditions. EVAs and emergency operations would be conducted in compartments or suits of 8 psi. These estimates are based on current STS conditions.

It is anticipated that food onboard a NASA space station will be somewhat better (variety, texture, etc.) than is currently available on STS missions. It must also be expected that improvements will be made in the areas of hygiene. We assume, however, that full body showering will remain a luxury.

Physical Quality of Habitat. Since the building blocks that will be used to construct a NASA space station must be shuttle-compatible (i.e., they must fit in the orbiters' cargo bays), it is assumed that Spacelab-type modules will be used. For this reason, we assume that the physical qualities of the station will be similar to those aboard the orbiters and Spacelab, although modified for long-duration occupancy.

Sells, S.B. The taxonomy of man in enclosed space. In J.E. Rasmussen (editor) **Man In Isolation and Confinement,** 1973, pp. 281-303.

*NOTE: This presentation summarizes interim results of research sponsored by the Manned-Vehicle Systems Division of the NASA/Ames Research Center (NAS2-11690). Conclusions drawn from the research are the author's responsibility and do not necessarily reflect NASA opinion or policy.

AD-P003 354

Criteria for Selecting Subjects for the Assessment of Advanced Crew System Concepts

Gilbert G. Kuperman
Air Force Aerospace Medical Research Laboratory

Abstract

A five-Laboratory working group was tasked with the assessment of crew system concepts being considered for the dual role fighter. The methodology employed was based on exposing rated TAC personnel to these concepts through the "role playing" of a highly structured mission scenario. Because many of the avionics subsystems of interest are developmental, the issue of subject experience was considered to be critical to the validity of judgements elicited in subject de-briefing.

Introduction

General

The human factors practitioner, particularly in a military research laboratory environment, is frequently confronted with the task of evaluating conceptual (i. e., pre-breadboard) approaches to the man-machine interface. Most often, this tasking must be carried out in a context in which the machine side of the relationship involves systems whose capabilities are based on developing technologies rather than on existing hardware. A variety of paradigms (e. g., Kearns, 1982) have been attempted to provide at least first order insight into the effectiveness of the design approach. In the Air Force, and with specific regard to the design evolution of aircraft cockpits, these approaches have included artists' concepts, scale paper cutouts, and a variety of scale and full sized crew station mockups. These approaches can provide an excellent environment within which the researcher can explore expert opinion as to the facilitating and limiting aspects of the conceptual crew station design.

Background

The Tactical Air Command (TAC) validated a Statement of Operational Need for a dual role fighter (DRF). This weapon system is to be a replacement for the F-4D and is to exhibit increased range and payload capabilities in both the air-to-air (A/A) and air-to-ground (A/G) arenas (hence the "dual role"

descriptor). A derivative version of either the F-15 or the F-16, equipped with suitable weapons, sensors, and avionics, is to provide the airframe. The weapon system is to be capable of through-weather navigation and under-weather target acquisition and weapon delivery, day or night. Ritchie (1983) provides a colorful description of the DRF mission.

As a preparatory step in the weapon system acquisition process, the Aeronautical Systems Division (ASD) requested that the Air Force Systems Command laboratories provide early assessments of the conceptual crew stations being developed for the F-15 and F-16 variants (Berry, 1983). This request was responded to through the formation of an *ad hoc* working group composed of representatives of the Air Force Avionics, Flight Dynamics, Armament, Human Resources, and Aerospace Medical Research Laboratories. Close coordination with the weapon system user, the Tactical Air Command (TAC), was accomplished through the participation of Headquarters TAC, local TAC Systems Office, and operational unit personnel throughout the course of the assessment effort. Communications between the working group and the aircraft developers (McDonnell Aircraft Company for the F-15 and General Dynamics/Ft. Worth for the F-16) were maintained through the ASD System Program Offices.

Crew System Issues

TAC required that the DRF be crewed by a pilot and a weapon system officer (WSO). The conventional F-15 and F-16 aircraft are single place. The ability of the crew system to support situation awareness and crew coordination were selected as assessment dimensions.

The DRF is to be equipped with a variety of advanced avionics (e. g., the LANTIRN terrain following radar, laser designator/ranger, missile boresight correlator, and target acquisition and navigation forward-looking infrared sensor pods) and advanced weapons (e. g., AMRAAM and precision-guided munitions) in order to be capable of accomplishing the dual role mission. Allocation of crew duties and crew workload were also selected as crew system assessment dimensions.

The DRF crew will have to overcome the high level of complexity inherent in applying a suite of advanced technologies under conditions of high threat, adverse weather, and darkness in carrying out missions which may combine both A/A and A/G tasking. The capabilities of the crew/system interface to support weapon system mission effectiveness was selected as a fifth dimension along which to perform the assessment.

Methodology

Kearns (1982) provides a detailed description of the use of static (i. e., non-interactive) crew system mock-ups in the evaluation of alternative or conceptual aircrew station designs. Basically, full-scale plywood stations are overlayed with

photographic renditions of all panel-mounted controls and displays. Other control heads (e. g., stick and throttle) are included as non-functioning models. Subjects (Ss) are walked through representative tasks at a level that includes switchology inputs and display responses. Kuperman, et al. (1983) provide a discussion of the application of this "role playing" familiarization process for the case of the DRF.

Selection Criteria

Rated Personnel

The Ss used in the DRF crew system assessment acted as a panel of experts providing the experimenters with informed opinions, anecdotes, and qualitative judgements regarding each design concept. The "role playing" and the "ground school" activities which prepared the Ss for participation were essentially highly structured approaches to concept orientation and control/display familiarization. Rated personnel were required in order to assure understanding of the mission, validity of approaches to avionics and weapons utilization, and for face validity in assessing potential acceptance of the crew system concepts by the using community (i. e., TAC).

Weapon System Experience

It was felt that the Ss should be drawn from rated TAC personnel who were current on one of four present-day aircraft: F-4, F-15, F-16, and F-111. F-4 personnel were desired for two reasons. First, since the DRF is intended as a replacement weapon system (albeit with greatly improved capabilities), F-4 crew members should be most familiar with the mission. Second, the F-4 is also a two place aircraft and Ss drawn from this population should be experienced in the areas of crew coordination and allocation of functions between crew members.

Since the DRF will be a variant of either the F-15 or F-16, personnel current on these single-place aircraft should be most familiar with the capabilities of these aircraft and with the A/A aspects (threats, tactics, weapons) of the DRF mission. Although two-place versions of these aircraft exist, they are pilot trainers and Ss drawn from the F-15/-16 communities are pilots and have not received the specialized training provided to WSOs.

The F-111 is also a two-place aircraft and provides Ss who are either pilots or WSOs and who are sensitive to the issues of intra-crew coordination and allocated duties. The overriding reason, however, for selecting Ss from the F-111 population was because this weapon system possesses a terrain following radar and, when equipped with the PAVE TACK forward-looking infrared (FLIR) pod, is the currently operational weapon system with capabilities very similar to those required for the DRF.

Sub-system Experience

The DRF will be equipped with an avionics suite equal to the complexities of its mission. Berry (1983) and Ritchie (1983) describe many of the avionics and other capabilities to be represented in a DRF weapon system. These required capabilities include:

- terrain following radar
- navigation and target acquistion forward-looking infrared sensors (FLIRs)
- wide field-of-view head-up display (HUD)
- multi-mode (A/A and A/G) radar
- missile boresight correlator
- target designation/ranging laser
- secure communications
- advanced fire control computer
- advanced navigation (map) display

Further, the DRF must be compatible with the accurate delivery of both gravity ("dumb") and precision-guided ("smart") weapons. Many of these hardware-supported capabilities are based on technologies only now undergoing development (e. g., the LANTIRN HUD). As has already been suggested, the F-111 and F-4 PAVE TACK configurations provide the closest approximations to these avionics capabilities in the present-day inventory.

Subjects

Twenty TAC personnel were actually employed in carrying out the "role playing" activity. Eleven of these Ss were rated pilots and nine were experienced WSOs. Role playing was always conducted with a pair of Ss (one pilot and one WSO) working together as a crew. The five pilots who participated in the F-15 DRF portion of the exercise had an average of approximately 2500 flying hours experience while the six F-16 pilot Ss averaged about 1900 hours. Two were current in the F-15, two in the F-16, two in the F-4, three in the F-111, and two in the F-5. The WSOs were all current in either the F-4 or the F-111. The average experience level of the WSOs was approximately 1500 hours (F-15 Ss) and 1080 hours (F-16 Ss), respectively.

Benefits of the Process

The use of carefully selected Ss in carrying out the crew system assessment process resulted in a very high level of face validity. Using command (TAC) acceptance of the results was strongly supported by the participation of using command personnel in attaining them. (In fact, the author was invited to participate in a TAC Cockpit Configuration Working Group which addressed "missionizing" the crew system.) Weapon system development organization (ASD) acceptance was demonstrated by the briefing of findings and recommendations back to the respective contractors prior to their submission of formal proposals for the DRF. Further, these findings and recommendations were included as part of the proposal evaluation process.

Conclusions

If Ss are selected on the bases of well thought out criteria, relatively simple methods of assessing the utility of advanced man-machine interface concepts can yield meaningful results. Basing such criteria on operational requirements for the subject weapon system helps to assure using command participation in the process and acceptance of its results. The early acquisition of expert opinion data as to the merits and liabilities of a proposed crew system can be used to guide the application of more sophisticated (and expensive) assessment tools. Spencer (1983), for example, describes the use of computer-based, highly interactive simulation facilities to the refinement of the F-15 DRF crew system.

References

Berry, F. Clinton, "The Revolutionary Evolution of the F-16XL," Air Force Magazine, November 1983.

Kearns, J. H., A Systems Approach for Crew Station Design and Evaluation, Final Technical Report, AFWAL-TR-81-3175, Air Force Wright Aeronautical Laboratories, Wright-Patterson AFB, Ohio, February 1982.

Kuperman, Gilbert G., Moss, Richard W., and Bondurant, Robert A., "Crew System Assessment Methods applied to Dervivative Fighter Cockpits," Proceedings of the Human Factors Society 27th Annual Meeting, Norfolk, Virginia, October 11-14, 1983.

Ritchie, Steve, "An Eagle for All Arenas," Air Force Magazine, November 1983.

Spencer, David C. "F-15 Dual Role Fighter Cockpit Integration," Proceedings of the Society of Automotive Engineers 1983 Aerospace Congress and Exposition, Long Beach California, October 3-6, 1983.

Relationship between Crewmember Characteristics and Tank Crew Performance

Bruce Sterling
United States Army Soldier Support Center

Abstract

Numerous prior studies have shown relationships between individual soldier characteristics (such as general and specific mental abilities, education, etc.) and individual performance. However, few studies have demonstrated a valid relationship between individual crewmember characteristics and crew or weapon system performance. This study examined relationships between individual characteristics of tank Commanders (N=166) and gunners (N=165) and performance in a crew drill - Table VIII gunnery. Results suggested that for gunners, mechanical ability and experience related to crew performance, while for tank commanders overall mental ability related to crew performance. Results could be used in assignment of soldiers to these crew positions.

Introduction

The question concerning what sort of people make the best fighters has been asked for at least several thousand years, as we shall see later. However, the first scientific (that is, data based) attempts to answer the question took place during World War II (Stouffer, Lumsdaine and Lumsdaine, 1949). Soldiers in infantry squads were divided into three categories of fighting effectiveness (below average, average, or above average) by individuals familiar with their behavior under fire. Soldiers for whom substantial agreement on fighting ability could not be reached were dropped from the analyses. Comparison of background characteristics of these groups showed that better fighters were older (25 or over), more likely to be married, had higher mechanical aptitude scores (top two of seven categories) and were more often at least high school graduates. A similar study differentiating fighters from nonfighters was performed during the Korean Conflict (Egbert, Meeland, Cline, Forgy, Satchler, and Brown, 1957). Results of extensive psychological tests and interviews with fighters and nonfighters showed that fighters were more intelligent, more masculine (e.g., enjoyed contact sports), doers (e.g., had more hobbies), more socially mature (e.g., took responsibility for others) and preferred by peers, more emotionally stable (e.g., had fewer anxiety symptoms), better leaders (e.g., had higher social status with their peers) and were physically healthier (e.g., less reported illnesses). They also reported a more stable and happy home life, had a greater fund of military knowledge and greater speed and accuracy in manual/physical tasks.

Individual characteristics (particularly intelligence) have also been shown to relate to soldier performance during peacetime. Vineberg and Taylor (1972), Armor (1981) and Shields and Grafton (1983) showed that over a variety of Military Occupational Specialties individuals in higher mental categories were more likely to have passed their Skill Qualification Test, a measure of individual job performance.

However, a growing number of today's weapons are crew served. Would similar factors relate to the combat effectiveness of crew members of the high-tech crew served weapon systems prevalent in modern combat? To answer this question, Wallace (1982) conducted a study comparing the mental aptitude of tank crewmembers with weapon system performance in the 1981 Canadian Cup international tank crew competition. This was called the Gideon study, after the individual who led 300 soldiers hand-picked by the Lord to an overwhelming victory over the entire Midianite Army (see the Book of Judges, chapters 7 and 8). Results indicated a rather large and reliable relationship between Armed Forced Qualification Test (AFQT) scores of tank commanders and points scored by the tank crew in the competition. However, this study has several serious flaws. For instance, crews in this competition were highly selected and thus probably not typical of armor crews in general. Also the number of crews involved was quite small (N=15) and two of these observations which did not fit the regression line were dropped. This study represents an attempt to use archieval data to examine the relationship between crewmember characteristics and weapon system performance.

Method

Subjects

Subjects were 166 U.S. Army tank commanders (TCs) and 165 gunners who took part in crew live fire exercises in Grafenwohr, Germany during mid-1982.

Instrument

The Weapons Crew Performance Information System was developed by the U.S. Army Training Support Center (ATSC). This form can be used to collect information (e.g., weather conditions, visibility, crewmembers' social security numbers, crew performance, etc.) on live fire exercises.

Procedure

The above form was used to collect data on TCs and gunners of crews undergoing table VIII qualification runs at Grafenwohr, Germany during mid-1982. Using the social security numbers, Soldier Support Center was able to obtain enlisted master file (EMF) data on 331 of the 509 (65%) TCs and gunners who completed the form. Using the two data bases (EMF and ATSC form), the following information was recorded for each TC or gunner: rank, AFQT percentile, mental category, combat subscale score, general technical subscale score, mechanical maintenance subscale score, education (high school graduate/non high school graduate), position (commander/gunner). Table VIII qualification results (distinguished, qualified on first round, qualified on second round, unqualified), months assigned with this crew, months since last live-fire exercise, visibility on range (good/poor) and whether they were a platoon sergeant/first sergeant or not.

Results

Analyses were performed separately for TCs (n=166) and gunners (n=165). Pearson correlations were performed between Table VIII qualification results and commander/gunner rank, AFQT percentile, combat subscale score, general technical subscale score, mechanical maintenance subscale score, months

assigned to crew and months since firing. Results, presented in Table 1, show for TCs, a reliable but small positive relationship between AFQT percentile rank and performance in Table VIII gunnery. The correlation between Table VIII performance and the mechanical maintenance subscale score was actually larger, but because of the smaller number of TCs with this score in their records, it did not achieve statistical significance. For gunners, small but reliable positive relationships existed between Table VIII performance and rank, mechanical maintenance subscale score and months with the crew.

Table 1

Pearson Correlations With
Table VIII Gunnery Qualification
Results

	Tank Commanders	Gunners
Rank	N= 166 r= .115 P= NS	N= 165 r= .191 $P < .05$
AFQT %ile	N= 164 r= .197 $P < .05$	N= 164 r= .003 P= NS
Combat	N= 50 r= .005 P= NS	N= 137 r= .148 P= NS
General Technical	N= 158 r= .149 P= NS	N= 161 r= .066 P= NS
Mechanical Maintenance	N= 51 r= .256 P= NS	N= 137 r= .168 $P < .05$
Months in Tank	N= 166 r= .137 P= NS	N= 165 r= .167 $P < .05$
Months Since Firing	N= 166 r= .095 P= NS	N= 165 r= .043 P= NS

Note: NS = not significant

Chi Squares were computed between Table VIII qualification results and TC/gunners' education, range visibility, platoon/company leadership position (yes or no) and mental category. While no statistically significant relationship emerged for either crew position, it appeared that there might

be a relationship between mental category and whether a TC's tank at least qualified on the first round or not. When performance on Table VIII was dichotomized between distinguished/first round qualified and second round qualified/unqualified, a statistically (and practically) significant relationship was found between Table VIII performance and mental category of TCs (see Table 2). While 100% of the Category I TCs qualified on the first round, only 56.5% of the Category IV TCs performed similarly.

Table 2

Relationship between TCs Mental Category and Table VIII Gunnery Qualification Results

	MC1	MC2	MC3A	MC3B	MC IV
Distinguished/ Qualified 1st Rnd	N=4 100%	N=46 86.8%	N=24 82.8%	N=46 83.6%	N=13 56.5%
Qualified 2nd Rnd/Unqualified	N=0 0%	N=7 13.2%	N=5 17.2%	N=9 16.4%	N=10 43.5%

Chi Square = 11.40
degrees of freedom = 4
$P < .05$

Any significant relationships with Table VIII qualification results are surprising considering the low amount of variance in the dependent variable. That is, overall 2.4% of the tanks were distinguished, 77.7% qualified on the first round, 14.8% qualified on the second round and 5.1% did not qualify. Thus with a performance measure which discriminated more among tanks, the relationships between these variables and crew performance could become much larger.

Discussion

Overall, results show that for tank commanders, overall mental ability, represented by AFQT percentile or mental category, was most related to crew performance. While the magnitude of the AFQT correlation was small, the mental category results showed that over 25% fewer category IV TCs qualified, compared to any other mental category. These results could attain practical significance for the Army in combat. For example, if it were true that only tank crews qualifying on the first round are combat effective, about four of five tanks with an MC I to III B TC would be combat effective while only about one of two tanks with an MC IV TC would be combat effective.

For gunners, rank (a surrogate measure of experience, maturity and many other things), the number of months spent with that particular crew, and mechanical maintenance ability related to Table VIII performance. Overall mental ability (i.e., AFQT score or mental category) did not relate reliably

to performance. These results differ from a recent Israeli study (Tziner and Eden, 1983) where overall mental ability of all tank crewmembers (save TC who was not studied) were related individually and collectively to superiors' ratings of crew performance.

This pattern of results suggests that mechanical ability and experience may influence gunnery, while overall information processing/decision making skills may be more critical for TCs. This makes sense given the nature of the two jobs.

These data suggest that instead of one position being more important than another in overall weapon system performance, each position may be important. However, different abilities/characteristics affect system performance at different crew positions. These results should guide the design of future studies and research in crew performance.

References

Armor, D. J. Mental Ability and Army Job Performance (Rand Draft Report). Santa Monica, CA: The Rand Corporation, 1981.

Egbert, R. L., Meeland, T., Cline, V., Forgg, E. W., Satchler, M. W., & Brown, C. Fighter I: An analysis of combat fighters and non-fighters. (HumRRO Technical Report 44). Presidio of Monterey, CA: US Army Leadership Human Research Unit, under technical supervision of George Washington University Human Resources Research Office, 1957.

Shields, J. L., & Grafton, F. C. A natural experiment: Analysis of an almost unselected Army population. Paper presented at the meeting of the American Psychological Association, Anaheim, CA, August 1983.

Stouffer, S. A., Lumsdaine, A. A., & Lumsdaine, M. H. Attitudes before combat and behavior in combat. In S. A. Stouffer, A. A. Lumsdaine, M. H. Lumsdaine, R. M. Williams, Jr., B. Smith, I. L. Janis, S. A. Stair, & L. S. Cottrell, Jr. The American soldier: Combat and its aftermath, Vol II of Studies in Social Psychology in World War II, Princeton, NJ: Princeton University Press, 1949.

Tziner, A., & Eden, D. Effects of tank crew composition on tank crew performance: Does the whole equal the sum of its parts? Paper presented at meeting of the American Psychological Association, Anaheim, CA, August 1983.

Vineberg, R., & Taylor, E. Effects of aptitude (AFQT), job experience, and literacy on job performance: Summary of HumRRO work units UTILITY and REALISTIC, TP-71-1. Alexandria, VA: Human Resources Research Organization, February 1971.

Wallace, J. R. The Gideon criterion: the effects of selection criteria on soldier capabilities and battle results, Research Memorandum 82-1. Fort Sheridan, IL: Research, Studies and Evaluation Division, Program Analysis and Evaluation Directorate, January 1982.

Field-Expedient Maintenance Experiences of M60-Series Tank Crewmen

Bob G. Witmer
US Army Research Institute-Fort Knox

Abstract

Although most tank crewman duties involve operating rather than maintaining their vehicles, the ability of crewmen to make expedient repairs in the field could be crucial to the crew's survival. Preliminary reports suggested that tank crewmen occasionally employ unauthorized field-expedient repairs during training exercises when a mechanic is unavailable. To determine the types and effectiveness of field-expedient techniques, 76 incidents of field-expedient maintenance were collected during interviews with 33 armor NCOs. The incidents were reviewed and grouped into eight categories representing different types of field-expedient techniques. Each of the eight categories represents a generalized maintenance approach that might be useful in a number of emergency situations where expedient repairs are essential.

Introduction

Standard Maintenance Procedures

Under normal circumstances, most troubleshooting and repair of complex military hardware such as tanks is performed by qualified mechanics and technicians. These mechanics and technicians, trained in approved troubleshooting and repair techniques, accomplish the repairs using procedures prescribed in detail in voluminous technical manuals. In theory at least, these mechanics and technicians make repairs by the book, using authorized parts and procedures to effect repairs in the required manner. In practice, however, technical manuals cannot possibly cover every problem that might occur with a complex weapon system; thus mechanics and technicians must depend to some extent on their troubleshooting skills to correctly diagnose system faults.

Siegel and Jensen (1955) have suggested that effective troubleshooting involves hypothesizing the cause of a malfunction from observed symptoms and testing the hypothesis by making diagnostic equipment performance checks. Not every malfunction, however, requires such careful troubleshooting. In some cases the cause of the malfunction will be obvious, even to a relatively inexperienced mechanic. Extensive experience in evaluating and repairing malfunctioning equipment often allows the experienced troubleshooter to go directly from the symptom to the repair without making any diagnostic checks. Such shortcuts may be effective, especially in those cases in which an observed symptom is almost always associated with a particular fault. On the other hand, shortcuts may lead even the most experienced technicians to misdiagnose the cause of a malfunction (Chalmers, 1957).

Field-Expedient Maintenance

Unlike trained mechanics and technicians, tank crewmen are given little formal training in troubleshooting and repair procedures. Crewmen are trained to perform routine checks and services and to refer all other malfunctions to organizational maintenance. This division of responsibility between tank crewmen and skilled mechanics works well enough in peacetime, but combat presents special problems. When their tank sustains damage or malfunctions during combat, and trained maintenance personnel are not readily available, crewmen may be forced to rely on their maintenance skills to extricate themselves from life threatening situations. Because of the immediacy with which the repairs must be accomplished and the lack of approved repair parts, crewmen in these situations may have to resort to the use of unauthorized materials and techniques for effecting the repairs.

While tank crewmen have long recognized the value of making a timely repair in the field, armor experts have only recently begun to see the advantages of field-expedient techniques. Some armor experts are now suggesting that senior NCOs and other leaders be trained to perform field-expedient techniques. But detailed information on field-expedient maintenance has been lacking. Only spotty informal reports of field expedient maintenance have been available.

Critical Incident Technique

Because performance of field-expedient maintenance does not occur with any predictable regularity, obtaining sufficient information about field-expedient maintenance through direct observation is not feasible. However observations of field-expedient maintenance can be gathered indirectly through a method known as the critical incident technique (Flanagan, 1954). This technique was developed by Flanagan and his colleagues for determining what behaviors were critical to effective and ineffective performance of job activities by Army Air Force aviators. The critical incident technique has since been used to determine the critical requirements for effective performance in a variety of different jobs (Fivars, 1973). The technique involves asking competent observers to describe incidents in which the behavior of an individual was particularly effective (or ineffective) in performing a prescribed activity or job. The critical incident technique is used in the present study to obtain information about the kinds of unauthorized procedures used by tank crews to repair their vehicles.

METHOD

Procedure

Thirty-seven armor NCO's were asked to describe incidents in which they, as members of tank crews, performed or saw others perform unauthorized maintenance techniques that were clearly effective in restoring a disabled or malfunctioning tank to operation. To ensure the inclusion of pertinent information, each NCO was asked a series of questions about the incidents described. Questions asked for: (1) the circumstances under which the maintenance occurred; (2) initial symptoms suggesting a malfunction; (3) troubleshooting checks made by the crew; (4) symptoms leading to fault diagnosis; (5) the

faulty system or component; (6) how the repair was made; (7) tools used in making the repair; and (8) how long the repair took. The NCOs provided the information anonymously during individual structured interviews conducted by interviewers from the Army Research Institute.

Analyses

The above procedure yielded two kinds of information: (1) field-expedient maintenance incidents based on the personal experiences of armor NCOs; and (2) profiles of the incidents, characterizing incidents along each of several dimensions. The field-expedient maintenance incidents were sorted into groups of similar incidents, which became the basis for eight distinct categories of field-expedient maintenance. Fifty of the 76 incidents were used to derive the categories initially; these categories were then used to classify the remaining 26 incidents. On the basis of the successes obtained and difficulties encountered in classifying the remaining incidents, the category definitions were refined. As a measure of the reliability of the categories, two persons not affiliated with the study independently classified the 76 incidents. The percentage of agreement among the classifiers was computed. The information contained in the incident profiles was evaluated by determining the proportion of incidents falling under the different levels of each dimension.

RESULTS & DISCUSSION

Descriptive Analyses

Descriptive analyses of the incident profiles revealed some interesting results. Approximately two-thirds of the field-expedient maintenance experiences occurred during collective training exercises, such as ARTEP's, field problems, and gunnery. In looking for the source of malfunctions, crewmen seldom reported making troubleshooting checks, other than a quick visual inspection of the suspected component. Only 24% of those interviewed reported using troubleshooting procedures to isolate the problem. When tools or supplies were required to make the repair, soldiers reported selecting from issued items (e.g., wrenches, sockets, track jacks, flashlight batteries) and non-issue items such as sticks, electrical tape, or a spring from a ball-point pen. Some soldiers carry special tools and supplies for the express purpose of making field-expedient repairs, such as vise grips, 90-mile-an-hour tape, a green sticky tape for repairing air hoses, and canned ether for starting a cold tank engine. Some repairs are made without using any tools or supplies whatsoever. For example, a soldier might manually operate a broken steering linkage or spit on the back of a round to increase conductivity so that it fires. Using available tools and supplies and their imagination, the soldiers completed the typical field-expedient repair in an hour or less.

Maintenance Categories

Table 1 shows the eight categories of field-expedient maintenance. While the categories listed in Table 1 were based entirely on the experiences of M60-series tank crewmen, the categories may apply to field-expedient maintenance on other systems as well. Due to the relatively small number of incidents collected, however, new categories may appear when maintenance of other

weapon systems is considered. Still, the categorization of field-expedient maintenance has identified various approaches that may be used to make field-expedient repairs in a wide variety of situations. For example, soldiers might be taught that when a faulty part or component is known to be interfering with carrying out their mission, and cannot be mended, then they should consider using a substitute part, bypassing the part, or removing the part completely and operating without it. Similarly, soldiers could be made aware of other approaches (e.g., mannual assist) that might be useful in certain kinds of situations.

The adequacy and reliability of the categories of field-expedient maintenance were determined by comparing the author's classification of the 76 incidents with that of each of two independent classifiers. The first classifier placed 72% of the incidents in the same category as the author, and the second classifier categorized 82% the same as the author. Working independently the two classifiers agreed on 72% of the incidents.

Table 2 contains one example from each category of field-expedient maintenance as reported by the armor NCOs. Incidents are listed for illustrative purposes only, and their

Table 1. Field-Expedient Maintenance Categories

Preventive Maintenance (Unauthorized) - Unauthorized maintenance performed to avoid anticipated problems.

Manual Assist - The soldier physically inserts himself as part of a malfunctioning system and manually assists the system as it operates.

Bypass/Remove w/o Replacement - A faulty component is taken out of the system and the system is operated without it, or the component is bypassed so that it no longer functions as part of the system.

Reposition or Adjust - A component that has become displaced, bent, jammed, locked, loose or out of adjustment is returned to its normal operating position.

Substitute Component or Part - A part or component is removed and replaced with an unauthorized substitute part.

Remove & Replace With Authorized Part - Either a component is removed and replaced that the crew is not authorized to remove or replace, or the manner in which the removal/replacement is accomplished (e.g., tools used, method used) does not follow accepted procedures.

Clean or Mend - A component is cleaned, patched, or mended by unauthorized personnel using approved methods or materials or by any personnel using unauthorized methods or materials.

Mechanical/Electrical/Chemical Assist - A vehicle is induced to operate or assisted in operation by applying an external mechanical, electrical, or chemical stimulus or boost.

Table 2. Examples of Field-Expedient Maintenance by Category

Preventive Maintenance - Due to vibration, the wedge bolt worked itself loose during an ARTEP. To prevent the wedge bolt from working loose again and eventually falling off, a hammer and chisel were used to notch the wedge bolt.

Manual Assist - During a field exercise at Fort Irwin (National Training Center), a tank was making a hasty attack across an open field laced with gulleys. The tank, moving at a fairly rapid clip, hit a deep gulley, causing the shifting linkage at the back of the engine to snap. The crew had to move the tank to avoid artillery shells that were being dropped behind them as they moved across the field. To move the tank, the TC got out on the back deck and, directing the driver's actions through the external phone system, manually operated the shifting linkage.

Bypass/Remove w/o Replacement - An M60A1 tank was on line preparing for an inspection at Fort Sill, Oklahoma. The start button was pushed, and nothing happened. Under the direction of a turret mechanic, one crewman used a wire to short across the starter relay, and the tank started.

Reposition or Adjust - The gunner was unable to adjust the brightness on the passive sight during a gunnery exercise. In examining the problem, the tank commander (TC) noticed that the whole reticle switch rotated when any attempt was made to adjust it. He knew immediately what was wrong with it. He took the plate off of the back of the switch and tightened a small nut that keeps the switch steady.

Substitute Component or Part - When the driver's seat does not move properly, the tank is normally deadlined. During an ARTEP a pin broke in the mount of the driver's seat where the handles are so that the seat would move neither up, down, forward, or backward. The tank commander substituted an Allen wrench for the broken pin and the seat worked perfectly.

Remove & Replace With Authorized Part - On a road march in Germany, during operations preventive maintenance checks and services suggested that the blower motors were defective. To return the tank to operation the malfunctioning blower motors were removed and replaced with good blower motors from a deadlined tank.

Clean or Mend - During tank gunnery exercises at Fort Polk a crewman on an M60A1 tank smelled a wire burning. The crewman visually checked for a burnt or broken wire. When the wire was located the soldier used WD1 (commo) wire to splice the broken ends back together. The exposed wire where the splice was made was then wrapped with tape.

Mechanical/Electrical/Chemical Assist - During cold weather in Germany, a tank would not start even when the crew tried to jump start it. To get it started, the TC injected "Start Pilot" (canned ether) into the air intakes. By using ether the TC was able to start the engine.

inclusion does not constitute a recommendation for their use. In the interest of brevity and clarity, the examples in Table 2 are paraphrased versions of the actual incidents. The incidents in Table 2 comprise only about 10% of those collected in this study, but even this small proportion demonstrates the ingenuity and creativity that tank crewmen exhibit in performing field-expedient maintenance.

REFERENCES

Chalmers, E. L., Jr. Evaluation of a method for teaching troubleshooting techniques. (Project AF 7790) Lowry AFB, Texas: Maintenance Laboratory, Dec 1957. (AD 842 255)

Fivars, G. The critical incident technique: A bibliography. (AIR MS. No. 890). Palo Alto, California: American Institutes for Research, 1973.

Flanagan, J. C. The critical incident technique. Psychological Bulletin, 1954, 51, 327-358.

Siegel, A. I. & Jensen, J. The development of a job sample troubleshooting performance examination. Journal of Applied Psychology, 1955, 39, 343-347.

Goals and Feedback in Command Group Training

Cal W. Downs
Kenneth M. Johnson
Kevin Barge
University of Kansas

Ira T. Kaplan
Army Research Institute
Ft. Leavenworth

Abstract

The use of feedback was investigated in five battalion command group training exercises provided by CATTS, the Combined Arms Tactical Training Simulator, at Fort Leavenworth in 1982-1983. Data were collected by interviews and questionnaires. A pre-training interview asked the battalion commander and his staff to describe both their unit and their personal goals for the exercise. A post-training interview explored (1) the degree to which unit and personal goals had been met and (2) the players' perceptions of the individual and group feedback that they had received. The results were analyzed with non-parametric statistics.

Some of the principal results were as follows: (1) There was great diversity in the goals of different command groups. (2) Within a group, however, the commander and his staff often had similar unit goals. (3) Reactions to feedback varied widely: although most welcomed it, some wanted less. (4) The feedback did not always address the players' goals. (5) Commanders who stated more goals evaluated the feedback as less adequate.

Introduction

Background

No concept is more basic today to theories of human communication than feedback, the loop by which information about actual performance is used to control future actions. Feeding information about performance back to individuals and interacting groups is also assumed to be a powerful means for enhancing organizational effectiveness, and most models of organizational effectiveness have a feedback component. Downs, et al. (1984) conducted a thorough review of the studies of feedback, and the predominant conclusion that could be substantiated by all of them was that feedback does indeed affect performance. The feedback process, however, is rather complex in that source, content, media, and receiver variables produce many different contingencies that affect the way individuals respond to feedback.

Feedback has obvious potential for maximizing effectiveness in training, and an opportunity arose to examine the feedback phenomenon in a military training setting. In 1982-1983 the Combined Arms Tactical Training Simulator (CATTS) at Fort Leavenworth introduced an expanded feedback component--an innovation which had not been present in previous training exercises--into the training of five battalion command groups. And in cooperation with the Army Research Institute and CATTS, the following field research was conducted.

Research Objectives

Most of the literature about performance appraisal prescribes some sort of goal identification, and Ivancevich and McMahon (1982) demonstrated that goal setting had a more pronounced effect on performance than a situation with no goals. From their review Downs, et al. (1984) concluded goals have a positive interaction with feedback in promoting performance. Therefore, this research focused primarily on the interaction between goals and feedback.

Goals

In an organizational setting, there is the possibility of having conflicting goals, and there is the question of who sets the goals. For example, previous research has failed to examine whether superiors' goals are internalized by subordinates. In the CATTS exercises, no goals had ever been set formally by the training groups; they were assumed to be present. Nevertheless, the nature of the training allowed for unit and individual goals to be set informally. Consequently, the research addressed the following research questions related to goals:

1. What kinds of goals were being set a) for the units? and b) for each individual position?

2. To what extent were the unit goals between the battalion commanders and their subordinates similar and different?

3. To what extent were the objectives generated by subordinates similar to those set by Battalion Commanders for their respective positions?

4. How did the similarities and differences in goals affect their responses to the feedback given?

Feedback

Two forms of feedback were given: 1) the entire command group shared their viewpoints in a group session, and b) each member of the command group was given feedback in a private session. Consequently, the following questions were framed:

5. How did the subjects evaluate the usefulness of the feedback?

6. How did reactions to the group and individual feedback sessions compare in terms of a) information adequacy, b) usefulness, c) completeness, and d) ease of understanding?

7. To what extent did the participants' number of a) unit and b) personal objectives and their similarity with those of the Battalion Commanders relate to their evaluations of the group and individual feedback sessions?

Participants

The participants were army and air force personnel in five battalion command groups which elected the CATTS training in 1982-1983. They participated in computerized battle simulations over a four day period. Each command group included a battalion commander (BC), S1, S2, S3, S4, air liason officer (ALO), and fire support officer (FSO).

Procedures

A pre-training interview occurred after the initial briefing but prior to the beginning of the exercise. Data collected from all but the battalion commanders focused on a) their individual goals, b) their unit goals, and c) the goals which they thought their battalion commanders had set for them as individuals. The battalion commanders, on the other hand, were asked to identify a) their goals for the unit, and b) their goals for each individual in the unit, including themselves. All of these interviews were taped with the permission of the participants.

Post-training interviews and questionnaires were administered after the completion of the training exercises. Data collected from the battalion commanders included the following: 1) changes in their unit training objectives, 2) changes in their personal training objectives, 3) perceptions regarding the adequacy of feedback information relative to all training objectives, 4) evaluations of the general usefulness of group and individual feedback sessions, and 5) preferences regarding the ways feedback could be given. The post-training interviews for the rest of the battalion staff collected the same data except for the changes in unit training objectives. All participants evaluated the group and individual feedback sessions using Likert-like scalar items, counterbalanced to control for ordering effects and alternately stated in positive and negative forms to control for random responses.

Because the data did not satisfy parametric assumptions, non-parametric statistics (Spearman Rho correlations, Mann-Whitney U tests, and Kruskal-Wallis tests) were used.

Results

1. There was great diversity in the goals of different command groups, and it was instructive merely to list these objectives while the simulation exercises remained the same. Both battalion commanders and battalion staff across the groups had very different kinds of goals. This, of course, makes planning to relate feedback to goals difficult.

2. Within a command group, however, the BC and his staff often had similar unit goals. Generally, less than half of the unit objectives stated by the BCs were also mentioned by their subordinates. (In the pre-training interviews, BCs averaged 5 unit objectives while their subordinates listed only 2-3.) However, of those training objectives listed by the subordinates, most of those goals were similar to those of the BC (Goal Similarity in Unit Objectives $\bar{x} = 2.0$). In general, the higher the number of unit objectives that subordinates could state, the more similar they were to those of the BC (Number of Unit Objectives with Goal Similarity in Unit, rho = .93). By command position, the subordinates most similar to and least dissimilar to the BC were the FSO, S1 and S4. The subordinates most dissimilar to the BC were the ALO, S3 and S2.

3. Within the superior-subordinate dyads, the commander and his individual subordinates had more similar goals than the subordinates anticipated. In terms of personal training objectives, most subordinates listed three, while their similarity to those of the battalion commander for the position was one or two objectives. In general, the higher the number of personal objectives set by the subordinate, the more dissimilar they tended to be from those set for them by the BC. On the other hand, the similarity was more than the subordinates anticipated. When asked to identify

what personal objectives their commanders might have set for them, the subordinate's similarity with their battalion commanders actually decreased (from $\bar{x}$ = 1.6 to $\bar{x}$ = 1.3 similar objectives).

4. Personal objectives tended not to change during the course of training. The correlation between Goal Similarity in Personal Objectives before and after the training was .92.

A major research focus was subject reaction to the feedback that they received. The commanders conducted group feedback for their staffs and then were given feedback from representatives of ARI who were studying various communication processes. Each staff member participated in the group feedback session and was also given individual feedback by the CATTS personnel. Some commanders gave individual staff members individual feedback, too. There was, of course, the presence of self-feedback, but it was not studied directly.

5. Individual reactions to the feedback sessions varied widely. Most subjects welcomed it, but there were some who wanted less of it. In general, however, all members of the command groups evaluated both group and individual feedback favorably. They rated the sessions on Likert-type scales of 1-7 with 1 = strong agreement and 7 = strong disagreement.

	Group Feedback	Individual Feedback
1. Agreed that the feedback was Useful	$\bar{x}$ = 2.26	$\bar{x}$ = 2.43
2. Agreed that it was Easy to Understand	$\bar{x}$ = 1.86	$\bar{x}$ = 1.46
3. Disagreed that the information was Incomplete	$\bar{x}$ = 5.74	$\bar{x}$ = 3.73
4. Disagreed that the information was Negative	$\bar{x}$ = 6.37	$\bar{x}$ = 4.33

These data indicate that the group sessions were viewed slightly more favorably than were the individual sessions. This may be explained in part by the comments made by many participants in which they expressed a desire for more detailed and more complete evaluation in the individual sessions.

6. Command group members felt that sufficient time was devoted to group feedback sessions and that there was sufficient involvement in them by CATTS personnel.

7. Four of the five battalion commanders would have liked more time in individual feedback sessions. Some even commented that they would have liked a much more detailed analysis than the overview often gives. Actually, the battalion commanders received no feedback specifically on their own individual performance. ARI representatives talked with them privately about general unit performance only; consequently, the only feedback received about their individual goals was self-feedback. With the exception of the S2, the other members of the command groups desired no more time in the individual feedback sessions.

8. Individual and group feedback sessions appeared to serve different functions. Group sessions focused on the unit objectives and the general training exercise. It was in the individual feedback sessions that personal goals and individual achievement was pursued.

9. Commanders who stated more goals evaluated the feedback as less adequate. In other words, there was a strong relationship between the number of goals set and perceptions of feedback adequacy. Normally, the battalion commanders would state four or five unit objectives. Of these, they reported receiving adequate information on three objectives and inadequate information on the rest. Some discovered that certain of their training objectives did not fit the CATTS program.

10. Similarly for command staff, the number of personal objectives was strongly related to their perceptions of feedback adequacy regarding those objectives (rho = .64). They averaged three goals and reported receiving adequate information on at least two of them ($\bar{x}$ = 2.23). Those personnel who reported receiving inadequate information about an objective tended to be the ones who had Goal Dissimilarity Regarding Personnel Objectives with their battalion commanders (rho = .32). These same people viewed the group feedback session as more incomplete than did the others. Conversely, those who reported a higher number of personal objectives on which they received adequate feedback evaluated the information given in the individual feedback more favorably.

Conclusions

The conclusions from this research must be tentative because of the limited numbers involved; it is hard to generalize from 5 groups. Nevertheless, this research can be seen as a first step in investigating two concepts that have immense potential for practical application in military training.

1. Participants felt that they benefitted from both the group and individual feed-back sessions. Because they served different purposes, they supplemented one another. The group session related to unit performance, and the individual session related to personal performance. ARI has also collected comparative data which demonstrates that these groups using feedback showed greater gains than did previous groups in CATTS which did not receive the specially designed feedback package.

2. While positive receptivity to feedback was not universal, comments about the weaknesses in the sessions were not an indictment of the desirability of feedback. Rather, they were expressions that the sessions could offer more. Frequently, comments in interviews suggested the feedback sessions should have been more specifically aimed at improvements. Battalion Commanders particularly desired more detailed, helpful feedback in individual conferences.

3. Goal setting was directly related to perceived adequacy of feedback. When subjects had feedback about their specific goals, they rated the feedback higher. Since formal goal setting was not the standard operating procedure for these training exercises, perhaps having the commander and subordinates talk about goals would enhance the training. These objectives could become targets at which to aim for performance improvement.

4. The results of this study indicate that the impact of goals and feedback on military training should be investigated on a wider scale.

Cal W. Downs, et. al., "A Review of the Literature on Feedback Research," Organizational Communication Abstracts, Sage, 1984. (in print)

Ivancevich, John M. and J. Timothy McMahon, "The Effects of Goal Setting, External Feedback, and Self-Generated Feedback on Outcome Variables: A Field Experiment," Academy of Management Journal, 25 (1982): 359-372.

An Examination of Air Force Supervisor-Subordinate Communication Relationships and Outcome Variables: A Review of the Findings[1]

Tim O. Peterson
United States Air Force Academy

Larry E. Penley
University of Texas, San Antonio

Brian L. Hawkins
Drexel University

Abstract

Over the last three years, the authors have examined a series of questions dealing with the communication relationships between supervisors and subordinates and some possible outcome variables in the military. The basic research question was to try to determine what types of supervisor communication lead to changes in performance, satisfaction and retention. In addition, the authors examined the effect of age and affiliation on the communication-outcome relationships. The paper will be a compilation of the findings over the last three years.

Introduction

There has been a great deal of interest recently in the relationships among communication, motivation, performance, and job satisfaction (Hawkins and Penley, 1978; Huseman, et al., 1978; Muchinsky, 1977b; Downs, 1977; Penley and Hawkins, 1979, 1980; O'Reilly and Roberts, 1977). The Air Force has also been interested in these relationships (Boyle and Krebs, 1979; TIG Briefs, 1977, 1979a, 1979b).

Although an increasing amount of research has been conducted in this area, few, if any, practical communication strategies have been developed. In order to develop practica[illegible] implications for this direction in communication research, it is important t[illegible]ine what kinds of communication are important, for what types of people. [illegible]'s exactly the type of contingency approach called for by Huseman, et al. [illegible] and Alexander and Penley (1981).

This paper is [illegible]port on our findings after developing a set of perceived communication behav[illegible] variables that could be related to outcome variables in an organizational setting.

---

[1]This is an edited version for the Proceedings. A more complete copy with references can be obtained from Major Peterson.

Finding 1

Performance was found to be higher when the supervisor was more receptive to listening to the subordinate. Performance was also found to be higher when the supervisor provided more information on how the employee was doing on his job. Both of these findings are consistent with prior research and are conceptually consistent with theory.

The finding that information from memos is negatively related to performance is harder to explain. The low performance associated with high receipt of memos may reflect an overload problem. Based on post hoc interviews, it appears that the organization uses written information to correct behavior. This policy is consistent with the negative sign. A discriminant analysis between those persons involved in high and low memo usage failed to show any demographic or organizational variables to explain this difference.

The findings concerning propensity to remain are interesting, but the non-findings may be more important. Only managerial receptiveness and responsiveness were significantly related to propensity to remain. The greater the extent to which an employee receives communication from upper management and perhaps feels part of the bigger picture, the more likely he or she is to stay with the organization. The fact that supervisory communication variables are not related to a person's propensity to remain is most interesting. A great amount of emphasis has been placed in the military on the supervisor's communication role in retention. These results, however, may suggest that the supervisor's impact is minimal. Although performance and satisfaction can be influenced by the supervisor's communication patterns, the intention to stay in the organization is evidently influenced more by the individual's personal characteristics. These findings tend to suggest that the overall managerial patterns or climate are extremely important and deserve more organizational attention to guarantee less turnover.

The results also indicate that older employees are more likely to stay in the organization. This makes sense in terms of the individual reducing dissonant feelings toward the job as he stays longer. The fact that older people are more likely to remain is also explained in terms of the lower potential of mobility of older employees.

The findings on job satisfaction are also consistent with prior theory. Those employees with the highest satisfaction are older, receive more performance information and management information, and their supervisors are more receptive. This scale reflects an affective response and hence is related to supervisory behaviors that indicate a personal interest in the employee.

Finding 2

The objective of the second analysis was to examine the relationship among five communication variables and the performance and job satisfaction as moderated by age difference between supervisors and subordinates. Age difference was computed to categorize the subordinate as a younger employee (more than five years younger), peer group employee (± five years of the superior's age), or older employee (more than five years older than superior).

Analysis of covariance was employed utilizing multiple regression techniques to control for the effects of ethnic difference and supervisor tenure.

Although a strong main effect exists among four of the five communication variables and job performance, these findings mask the results found when age difference is used as a moderator variable. In three of the cases, including the one insignificant main effect, age difference did not significantly contribute to the regression model. However, in two of the regressions, the moderation by age difference indicated that there was a significant relationship to performance only for older subordinates paired with a younger supervisor. The relatively flat slopes of the communication-performance relationship for the younger and peer group subordinates indicate that regardless of the supervisor's communication effort, performance will be essentially unaffected.

The steep slope, however, for the relationships between policy and personal communication with performance for older subordinates paired with a younger supervisor would indicate that there is very real potential for the manager to affect performance of this group by increasing his or her communicative efforts.

Turning to the relationship between these communication variables and employee satisfaction, the results are even more masked without using age difference as a moderator. There is a significant main effect for all five communication variables with job satisfaction. However, just as with the performance results, this masks the impact of difference in age. The relationship between satisfaction and the communication variables was extremely strong for younger subordinates. For younger subordinates, communication from an older supervisor directly and strongly impacted the employee's satisfaction. However, there was virtually no relationship between personal and career communication with satisfaction for the other two employee groups.

The same pattern of moderation also was found with task and policy communication. Although the relationship with satisfaction for peer and older subordinates was positive, it was significantly more positive for the younger subordinates paired with an older supervisor.

These results indicate that there is a differential impact of communication on outcome variables, depending on the age difference of the superior and subordinate. These results seem to indicate that communication most strongly affects satisfaction for employees who are younger than their superiors, but it tends not to impact their performance. The reverse is true of subordinates who are older than their superiors. In this case, communication efforts of younger supervisors impact their older subordinates' performance, but it has much less impact on satisfaction. Communication did not impact peer group subordinates to any significant degree.

Finding 3

The third research question was to determine if affiliation (military or civilian) of the supervisor and subordinate has an effect on the relationship between communication and propensity to remain (PTR). In this research, supervisor-subordinate pairs of affiliation were computed (e.g., CIV-MIL, CIV-CIV, MIL-MIL, MIL-CIV). In the analysis, each supervisor-subordinate combination was treated as a categorical variable by using dummy variables. Propensity to remain was treated as the dependent variable.

Task Communication indicates that there is significantly greater propensity to remain for both of the civilian-supervised groups when compared to the military supervisor-military subordinate group. Similarly, the civilian-civilian employee category was found to have higher propensity to remain when compared with a military supervisor and civilian employee. The relationships of task communication to performance was strongly moderated by supervisor-subordinate pairings. There was a significant and positive relationship between more communication on the task and an employee's propensity to remain, when the supervisor was a military supervisor. There was virtually no relationship between task communication and PTR when the supervisor was a civilian employee.

Policy Communication indicates the same basic pattern of results as Task Communication. The slopes of the relationships between Policy Communication and PTR showed strong, positive association when there was a military supervisor. Again the pattern of employees having overall better Propensity to Remain when supervised by civilian supervisors was found, with even greater statistical differences.

While the other three communication variables (Personal, Career, Managerial) showed a main effect with Propensity to Remain, they did not show a significant moderating effect when the supervisor-subordinate pairings were added to the regression.

Conclusion

The findings in this paper represent another step in showing the importance of the relationship between communication and other organizational variables. The findings point to the differences from one outcome variable to another in terms of the communication variables which are related to them.

In addition, the research shows the importance of examining simultaneously the source, message, and receiver characteristics in examining the relationship between communication variables and the Propensity to Remain variable. Finally, the findings indicate the influence that age difference plays on the relationship between communication and job satisfaction and job performance.

A further step in this line of research is to continue efforts to validate the communication indices which were developed as a part of this study. Efforts should also be made to determine whether communication variables are moderated by other demographic and organizational variables.

PANEL SESSION

PANEL SESSION: DEFINING THE CRITICAL PARAMETERS OF COMPUTER LITERACY

SESSION CHAIR: George E. Uhlig (University of South Alabama)

PARTICIPANTS: George E. Uhlig (University of South Alabama)

Russell N. Cassel (Cassel Psych Center)

Robert S. Block (TELEASE Corporation)

Robert S. Feingold (Headquarters Space Command)

Verne A. Weber, Jr. (San Diego State University)

PROCEEDINGS ENTRIES

"Computer literacy, information technology and national security" (George E. Uhlig)

"Defining the critical parameters of computer literacy" (Russell N. Cassel)

"The information utility and computer literacy" (Robert S. Block)

"A summary of U. S. Air Force computer literacy needs" (Robert S. Feingold)

"The situational computer literate" (Verne A. Weber, Jr.)

DEFINING THE CRITICAL PARAMETERS OF COMPUTER LITERACY (LTRCY)

Panel Members

George E. Uhlig, Dean (Chairperson)
College of Education,
UNIVERSITY OF SOUTH ALABAMA,
Mobile, Alabama 36688

Robert S. Block, President
TELEASE CORPORATION,
1875 Century Park East, Suite 930,
Los Angeles, California 90067

Lt.Col. Robert Feingold
HQ. UNITED STATE AIR FORCE SPACE COMMAND,
KRP-Peterson AFB, Colorado 80919

Major Verne Weber
AIR FORCE ROTC DETACHMENT,
San Diego State University, California 92182

Russell N. Cassel, Colonel, Air Force Retired.
Director, CASSEL PSYCH CENTER,
717 3rd Avenue,
Chula Vista, California 92010

George E. Uhlig

COMPUTER LITERACY, INFORMATION TECHNOLOGY AND NATIONAL SECURITY

The microcomputer, 59 million of which will have been sold and installed in the United States by 1989, is only the tip of a very large technological iceberg. This metaphor can be expanded if that iceberg is considered to be floating in a giant ocean, growing, ocean of information. It is generally agreed that we have now left the industrial and agriculture economy based society, and are in the information and service based one.

Information access is a severely limiting factor to world development, as well as to national security. New technologies are available which provide high-powered satellite delivery of information at a rate which in a 24-hour period is the equivalent of one million books. Receiver technology using a standard television set could cost less than $500 per household. Microcomputers, as they exist in several million homes, could be used for storage and data or word processing of received information.

At a time when the nation's schools are criticized for falling test scores, increasing attrition rates, and the inefficient use of public funds, many homes are becoming "learning centers" where youngsters are enriching and complementing school-based instruction with personalized learning experiences which are centered around the personal computer,

modem, and telephone line. In the view of many, youngsters having access to this sophisticated technology is a two-edged sword; many are becoming extremely well prepared for an affluent future in the field of electronics and computers; while a highly visible few are vandalizing national data structures. However, highly sophisticated strategies for security of data bases and transmissions systems are available.

Many believe that the national security of the United States will be highly dependent upon what the students and members of the military of today and tomorrow understand about technology, and specifically microcomputers. There is a real need for a curriculum of national security which includes specific competencies in R.O.T.C. programs at all levels and for students in all schools.

DEFINIING THE CRITICAL PARAMETERS OF COMPUTER LITERACY

Russell N. Cassel

It has been reliably estimated that some fair degree of computer literacy is essential for reasonable success in maybe as many as 75 percent of all job careers today. The question of what constitutes effective computer literacy is being debated with great fervor, but there is little agreement as to what should or must be included in the concept. Now in the 3rd generation of the computer, with the lazer threatening to become a fourth generation, the debate continues.

Placing Computer in Technology

Thomas Clayton, as Acting Director of Educational Technology, maintains that modern technology is but one life-time old, having begun in 1900. Civilization is maybe 800 life times old, and with 650 of those lifetimes being spent in a cave. About every 10 years, since 1900, there has been a major break-through in modern technology: 1900 Henry Ford made every man a neighbor; in 1910 the Wright Brothers gave us the airplane; in 1920 Marconi gave us the radio; in 1930 Warner Brothrs gave us motion pictures;in 1940 Hitler gave us the missile; in 1950 television came of age; in 1960 came the computer; in 1970 came systems analysis. The age of the computer had arrived.

Systems Analysis

The basis principles of systems analysis have served both the military and industry alike for many years in identifying critical elements essential for achieving well articulated goals,i.e., B-52 Weapons System, etc. The first and most basic principle involved in systems analysis is a careful articulation of goals. The second principle involves designing techniques for assessing progress towards those goals. The third principle seeks to identify available alternatives. The fourth and fifth principles deal with hazards and consequences for the

alternatives. By use of these principles we are able to establish the critical parameters for computer literacy, and nothing short of this will suffice.

Computer Literacy Goals

Every weapon system, every venture into space, and nearly every aspect of one's daily life invioIves computer based functions. Each day it seems these functions are becoming more and more sophisticated,i.e., the first generation robot could only sense things by feel, but the seond one could sense them by sight. It is obvious for us to achieve only modest goals in this regards, computer literacy must be comprehensive in nature, and must be evolutionary in perspective.

Cassel Computer Literacy Test

Parallel forms of the Cassel Computer Literacy Test have been developed for purpose of assessing computer literacy. The six part scores suggest the general content of this instrument: I. Computer Development, II. Technical Understanding, III. Computer Structure, IV. Information Processing, V. Information Retrieval, and VI. Communication System. Also, TOTAL SCORE, as overall computer literacy. The parallel forms permit assessment both before and after individuals are exposed to select computer learning.

Major Factors in Computer Literacy

A computer literacy workshop developed at The University of South Alabama for a Friday evening and all day Saturday depicts clearly the major factors involved in computer literacy. There are five independently organized modules used in this workshop, each containing one of the major factors presumed to be involved in computer luteracy: MODULE I - Computer Evolution and Historical Development, MODULE II - Computer Structure, Jobs, and Languages; MODULE III - Information Processing and Statistical Analysis; MODULE IV - Data Base Management Systems and Memory Retrieval; and MODULE V - Word Processing and Communication.

THE INFORMATION UTILITY AND COMPUTER LITERACY

Robert S. Block
President,
Telease, Inc.

In the mid-1970's, the United States entered the "Information Age." Shortly thereafter, the other industrialized nations followed. Since then, information, information technology, and communications technology growth have been accelerating at an increasing rate. This exponential growth is also increasing the information gap--a phenomenon most discussed in relation to developing countries, but one which is significant even within the most developed nations. Computer literacy is one aspect of this information gap.

Awareness of information has always been a major factor in the development of individuals, organizations, and societies. Increased access to information inevitably leads to increased awareness and thus to an improvement in performance. For all sectors of society, government, the public, and business, the need to know will grow.

When we consider technology trends, the growing need to access a growing volume of information, and the efficiency of electronic distribution of information vs hard-copy distribution, it is easy to envision the establishment of a satellite-based Information Utility. Such a Utility could be initiated in this country within a few years and by the end of the decade could spread worldwide. The Information Utility would provide a vast library of information available at almost every point on the earth's surface. The amount of information that could be delivered in a 24-hour period by high powered satellites would be the equivalent of millions of books each 24 hours. Hundreds of thousands of megabites of information would be available everywhere, every day. Indeed, the amount of information that could be handled by the Utility is virtually unlimited. Receiving antennas would be 2 to 4 feet in diameter, depending on location. Equipment capable of receiving, decoding, and displaying data, interactive displays (similar to video game technology), or full color motion on standard TV sets will cost less than $500, based on current technology. In the next few years, this price may include storage and data processing means as well.

The Information Utility is likely to begin as a one-way system which schedules transmission of information, based on its nature, the number of subscribers, and other pertinent factors. Some information may be transmitted continuously with loop time measured in seconds or fractions of a second. Other information would be available in minutes, hours, days, or weeks. From a user point of view, even in this one-way mode the system would have interactive characteristics. Receiving equipment could be programmed to automatically capture and store specific blocks of information at the scheduled transmission time for later user access and interaction. Storage requirements would be manageable, since storage would be necessary only for information to be used between scheduled transmissions. For very large subscriber and data base systems, this down-loading approach continues to appear to be more useful than on-line inquiry, even when two-way implementation becomes practical.

The Information Utility will probably resemble a common carrier, requiring data base suppliers to provide information in standardized formats at the appropriate time ready for encryption and transmission. Periodically (probably monthly) a unique decryption key will be transmitted addressed to each subscriber terminal. The key will permit decryption of specific information services. Based on usage, access time and data base fees will be securely accumulated in the subscriber terminal. The accumulated monthly charges will be displayed or printed for the subscriber. Timely payment of these fees will be required for future access to fee-based information.

Most of the elements necessary for the successful launch of the Information Utility are already available. Telephone-based data networks, teletext and videotext data and graphic systems, pay television networks, personal computers, videogame technology, secure encryption, and electronic billing are all available today. The high powered multi-channel satellite technology needed for direct broadcast satellite (DBS) will be available by 1986. The unused capacity of currently operating low and medium powered satellites is far more than the Information Utility would need. Using these satellites will require larger receiving antenna (about 7 feet) or terrestrial retransmission through cable or by multi-channel MDS systems. Initiation of the Information Utility requires the structured organization and integration of these elements and the subscribers to use it. Once implemented, the Utility would give each user enormous capacity to access information.

A fully integrated system such as the Information Utility will not only permit the use of computer data and interactive graphics, but also full motion color, sound, and perhaps information for other senses as well. The big job will be to train people to access and use that capacity. That should be the goal of computer literacy.

###

A Summary of U.S. Air Force
Computer Literary Needs

Robert S. Feingold - Headquarters Space Command

The U.S. Air Force is perhaps the most technically oriented of all the services. Air Force personnel daily operate and maintain complex equipment and weapons systems. Of necessity, Air Force personnel come in close contact with the most modern state-of-the-art technologies and concepts. Integrating these technologies and concepts into effective fighting systems and efficient support operations requires that Air Force personnel be educated and trained to quickly adapt to new situations where their systems will be employed.

Computers play a central role in modern weapon system operation, maintenance and unit support. In addition to the well known impact computer technology is having on weapons system design, computer technology has become a key factor in the design of effective combat support systems. For example, the Air Force supply system is computer based and provides direct support to fighting units, not only at their home base but while these units are deployed. Computers utilized in support of the supply function help reduce the manpower required by this functional area, reduce the costs associated with maintaining the large pipeline of replacement parts, and, most importantly, reduce the time required to effect weapon system repair thus enhancing readiness.

Many other support areas rely on computer technology to enhance mission support. But, in no case is computer technology a guarantee for mission success without trained personnel to employ it. Accordingly, the Air Force invests a great deal of time and energy in training personnel to effectively use computers properly. This training is directed toward achieving computer literate personnel who can effectively develop and maintain computer based systems and those personnel who use computer based systems for the mission. Not only do Air Force technical training centers provide basic skill training but comprehensive programs of on the job training provide unit unique training on a continuing basis.

The Air Force has recognized that computer technology is one important key to mission success and has committed the necessary resources to assure that Air Force people are comfortable with computers in their work environment.

The Situational Computer Literate

by

Verne A. Weber, Jr.

The traditional meaning of literacy is the ability to read and write. What is literacy to a laborer may be heresy to a grammarian, however. The criteria depend on the situation.

The concept of computer literacy is a function of the circumstances, too. At the most fundamental level, a computer literate must understand frequently-used jargon, be aware of what computers can do, and be able to log on and run simple programs. He should know how to retrieve, manipulate, and save information.

The literate does not necessarily need to know how a computer works, any more than an auto driver need be a mechanic. The start of being comfortable with a computer is any active participation with one, perhaps merely playing a video game. What is important is realizing that a computer is not an ogre, but just an incredibly useful and terribly stupid tool. Most people can leave comprehension of the innards of a computer to the architects and bit-fiddlers. User-friendly interfaces and (for programmers) high level languages have seen to that, creating Pascal users and other "quiche eaters."

After the literate has obtained a core of knowledge, he can sharpen his skills to fit his situation. He should look for ways to apply a computer in overcoming tedious tasks and improving productivity.

For the average user, this means finding out what is available in the way of hardware and software for a home or small office computer. He would need the ability to read consumer-oriented articles and carry on a conversation with a (down-to-earth) computer salesman. Financial management, word processing, reminders of things to do, recipe files, and whatever else he can imagine are only a keyboard (or mouse or joystick) away.

For the professional user, literacy is somewhat more complicated. A person who expects desired results from a systems analyst must be able to write specifications for what he wants, participate in design sessions, monitor progress, read the documentation, and test the products.

Finally, for the computer designer or the programmer, literacy may imply minimum competency in his craft. Writing code is not enough to keep a programmer's supervisor happy. Awareness of issues like run-time, memory usage, hardware restrictions, and support software availability are crucial. Debugging and testing skills and careful documenting must be second nature.

Thus, all who claim to be computer literates should possess a common core of knowledge about computers. However, the situation dictates what computer literacy is for the individual.

Terrain Travel Simulation: Data and Application

Nancy Mitchell, Army Research Institute
Robert Kraft, Grinnell College
Anne Martin, Decisions and Designs, Inc.

Abstract

Research in visual representation of terrain supported the development of images for simulated travel. Results showed that observers can estimate distances in photographs, both from the camera station to a target and between targets; this performance is analogous to distance estimation in the real world. Distance estimation performance is affected by the focal length of the camera lens at which the picture is taken. Visual coherence (the appearance of travel) can be maintained when photographs are taken at steps greater than the ten feet used previously.

Introduction

Background

The Army travels and fights primarily on land; therefore, the ability to visually simulate travel over land is of critical importance in both the system development and training arenas. If terrain simulation were available, human-machine interface and training issues could be addressed both earlier and more frequently in the development of a system which is dependent upon visually guided behavior such as the ability to move or guide a system over land. In order to address these needs, the Army Research Institute (ARI), in cooperation with Decisions and Designs, Inc. (DDI), has developed an interactive visual display of terrain at eye-level under a project titled, Advanced Terrain Representation (ATR).

ATR is an interactive visual display which simulates travel over open terrain. Interactive connotes control by the user; simulated travel refers to the ability to travel visually; and "over open terrain" means that visual travel is not confined to roads but would allow the user the same mobility available to a tracked vehicle.

In earlier work (Lippman, 1980), still photographs were used to generate videodisc based surrogate travel. The user had control over CRT images which moved in response to the user's command which was transmitted to a microcomputer. The microcomputer sent a message to the laser reading the videodisc to "skip a groove" and to go to the appropriate picture. It is because of the videodisc's variable presentation capability (skips pictures) and the computer control, that the travel simulation was completely interactive.

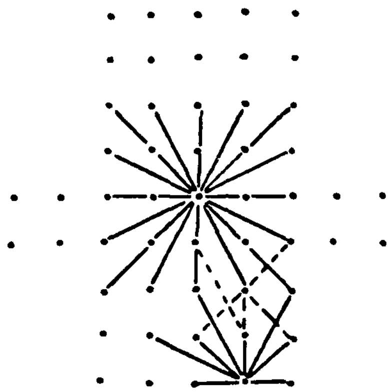
Lippman's surrogate travel constrained users to travel over streets; forward, backward, right and left. The objective of ATR was to extend this technology so that the system would emulate the ability of a tracked vehicle

to travel independently of roads, over open terrain. While the resulting visual travel is not as smooth as a moving picture, the effects of user control seem to be adequately compensatory.

Research Issues

The ATR project had two general objectives: the resolution of the logical problems involved in extending surrogate travel to a less restricted form of movement; and the determination of the optimal visual (spatial and temporal) resolution for producing an impression of motion.

The logical problems were resolved by utilizing a grid as shown in Figure 1. Each point on the grid represents a camera location. At each location, 16 photographs were taken with the direction of view being represented by the solid lines. Each direction of view is also a travel direction, and the broken lines show that travel can take place between any two adjacent points on the grid.



Computer control mandated that all photographs be taken with the camera on a level plane. Pilot research showed that visual travel up a hill was unrealistic under these circumstances if the hilltop was not in view. Increasing the camera lens angle provided the appropriate compensation but, at the same time, made objects in a scene appear more distant than they would appear if the focal length of the lens were set at "normal." The need to determine the effect of the camera lens angle (lens focal length) on the ability of humans to judge distances from observer to target and between two targets within a scene had critical implications for formatting the visual material to be used in this project.

Surrogate travel in Aspen was based on photographs taken at ten foot intervals. However, greater distances between photographs might still produce the desired visual coherence, the appearance of travel. If the photographs are too far apart, however, a viewer would become lost while traveling in a straight line because objects that should appear from scene to scene either change inappropriately or disappear altogether. The two experiments reported here were done to guide the development of ATR relative to the proper camera lens angle and the step size between pictures (visual coherence).

Experiment I

The perception of distance from the camera station to a target in a photograph can vary as a function of the lens angle at which the photograph was

taken. Camera lens angle, or angle of view (AOV), is a direct function of the lens focal length which ranges from telephoto (90mm, 30° AOV) to wide angle (35mm, 60° AOV). As the angle of view widens, objects appear smaller and farther away (Mascelli, 1965; Coynik, 1974; Giannetti, 1976).

There is a surfeit of psychological literature on how distance is perceived but very little information on real-world perception of distances over the range of interest to this particular project, from 10m to 1000m. The data which does exist (Gibson and Bergman, 1954; Gibson, Bergman, and Purdy, 1955; Teghtsoonian and Teghtsoonian, 1970) indicates that subjects increasingly understimate the distance to a target as that distance increases. When applying a power function to these data and those of Widen (1973a; 1973b), the exponents range from .67 to .86.

The purpose of this experiment was the determination of the viewing angle which most nearly induces the distance estimations produced in the real world. It was hypothesized that viewing angle would significantly affect the perceived distance tetween viewer and object in the manner specified by the aesthetic film literature. It was also hypothesized that the perceived distance between objects in the scene would remain unaffected by changes in viewing angle.

Method

Subjects. Twenty-four college students were paid five dollars an hour to serve as subjects. Subjects were run in five groups of two to six, and there were approximately equal numbers of males and females.

Stimuli and Apparatus. A total of forty standard 35mm color slides were taken: four apiece at ten locations. Four of the locations were in lightly wooded terrain, four in open terrain. At each of the ten locations, the four separate shots differed with respect only to focal length of the camera lens. The four focal lengths and the corresponding angles of view were as follows: 48mm, 48°; 28mm, 72°; 24mm, 84°; and 17mm, 104°. Each slide contained three naturally occurring target objects, one at each of three distances; near, mid, and far. For lightly wooded terrain, near was 0-50m, mid was 51-100m, far was 101-250m. For open terrain, near = 0-150m, mid = 151-450m, far = 451-1000m. The slides were arranged in four sets of ten with each of the ten locations represented once in each set. There was a practice set of two slides, and each set of slides was preceded by a title slide. The slides were projected onto a screen with a Kodak Carousel slide projector. Subjects were seated from 8 to 12 feet away and recorded their responses on a prepared data sheet.

Procedure and Design. Each subject viewed forty slides. While each slide was in view, the experimenter pointed out the target whose distance was to be judged by the subjects. Subjects were told to put themselves "in the scene as if taking the picture" and to judge the distance from themselves to the target. In addition, the slides of lightly wooded terrain contained objects separated by some distance that the subjects were to judge as well. A new slide was presented only after all subjects had responded to the previous slide. The slides were presented in one of four orders: 1, 2, 3, 4; 2, 3, 4, 1 etc. and all subjects responded to all slides. The design was a 2 (terrain) by 4 (viewing angle) within subjects design with three levels of object distance nested within terrain.

Results and Discussion

There was a highly significant effect of viewing angle for both lightly wooded terrain: [$F(3,69)=53.07$, $p < .0001$, MSe=2116.1]; and open terrain [$F(3,69)=11.37$, $p < .0001$, MSe=109,048]. As viewing angle widened, judgments of distance from observer to target increased accordingly. The results supported the notions of film theorists concerning the function of viewing angle.

The effect of differences in distances, i.e., near to far target, was also significant. For lightly wooded terrain, main effects were [$F(2,46)=104.51$, $p < .001$, MSe=16,130.5]; for open terrain [$F(2,46)=27.31$, $p < .0001$, MSe=476,607]. The interaction between distance and viewing angle was also significant for both types of terrain, $p < .0001$ in both cases. Finally, there was no significant effect for order of presentation for either type of terrain.

Table 1 is a summary of the average judged means over three targets per picture compared to the actual mean for both types of terrain at each lens angle.

Table 1

Judged Distance for Terrain Type and Viewing Angle

Terrain	Viewing Angle			
	45°	72°	84°	104°
Lightly Wooded				
near (0-50m)				
mean judgment	17.86	26.95	29.02	38.77
mean actual	32.25	32.25	32.25	32.25
mid (51-100m)				
mean judgment	41.00	56.20	64.13	83.49
mean actual	69.50	69.50	69.50	69.50
far (101-250m)				
mean judgment	115.95	147.71	161.21	196.93
mean actual	184.50	185.50	184.50	184.50
Open Terrain				
near (0-150m)				
mean judgment	31.43	42.72	50.86	80.66
mean actual	48.00	48.00	48.00	48.00
mid (151-450m)				
mean judgment	105.05	147.97	163.56	254.43
mean actual	205.00	205.00	205.00	205.00
far (451-1000m)				
mean judgment	276.30	369.95	457.29	544.98
mean actual	6.3.75	613.75	613.75	613.75

When the Steven's power function was applied to these data, the exponents were .83, 45° viewing angle, .84 for 72° VA, .89 for 84° VA, and .79 for 104° VA. The exponents for distance judgments on actual terrain ranged from .67 to .86.

Human observers can judge distance in photographs in the same manner as they judge distances on real terrain. While viewing angle does have an effect, the range of exponents for photographs was more narrow than the range of exponents for actual terrain judgments. This is probably the result of having only subject and viewing angle variance. The actual distance estimations exponents were compiled across experiments and experimenters, as well.

72 to 84 degrees is the more desirable viewing angle, since the mean judgments were closer to the actual means while still being underestimations, a distinguishing characteristic of human performance on real terrain. The judgment of distances between target within the photograph was unaffected by the camera lens angle.

Experiment II

A terrain based travel simulation must be visually coherent if it is to have an applied or research value. Visual coherence means that what the user is seeing "sticks together" in some orderly way, the display makes travel sense to the user. In the case of ATR, visual coherence is a function of step size on the grid (distance between photographs) and directions of view and travel (number of photographs) at any grid point (see Figure 1). The results from the experiment on angular displacement are outside the scope of this paper, however, the minimum number of 16 directions of view were used in the demonstration.

Method

Subjects. Twenty-four university students were paid five dollars an hour to serve as subjects. Subjects were run in five groups of three to eight and there were equal numbers of males and females.

Stimuli and Apparatus. The stimuli were eight six-shot sequences of 35mm slides. The first four shots were taken at 10m increments. The distance between slides four and five ranged from 10-40m for lightly wooded terrain and from 15-75m for open terrain. Distance between shots five and six duplicated the distance between shots four and five. In addition, there were eight sequences with random displacements between shots four and five. The sixth shot in the random displacement sequence was systematically related to the fifth shot. There were also four practice sequences with feedback.

Two Kodak Carousel slide projectors were connected to a Model 2 Kodak Carousel dissolver in order to present the sequences with no blank time on the screen. The effect was a rough approximation of movement from slide to slide.

Procedure and Design. After seeing each sequence, all subjects used a special answer sheet with 10m hash marks to indicate their perception of type and amount of displacement between shots four and five, and then between shots five and six. There was also a "don't know" option if they saw no relation between shots four, five, or six.

Data were analyzed separately for terrain type with first and second displacement judgments analyzed separately and pooled. Since lightly wooded terrain and open terrain could be formatted differently within the final system, there was no need to pool data across terrain type.

Results and Discussion

Data were analysed on the basis of a frequency count of judgments for linear, random, or angular displacement. A chi-square analysis for lightly wooded terrain yielded a significant effect of displacement level, [$x^2(6)=36.4$, $p < .001$]. For open terrain, displacement distance was significant [$x^2(6)= 12.4$, $p < .05$].

In lightly wooded terrain, the perception of coherent travel began to fall apart at 30m displacements and became markedly worse at a 40m jump. Jump size in open terrain was coherent up to and including 35m. The 75m jump actually produced a greater number of responses indicating coherence than the 55m displacements, but this could well have been an artifact of the number of distinctive objects in the display. This number was not controlled by count in the creation of the stimuli.

Conclusions

The results of these two experiments have shown us that the parameters for developing a system such as ATR are wider than was initially assumed. ATR was implemented using 12m steps between grid centers. This creates a step of 27m on a diagonal and the conjunction of these step sizes may be disconcerting and require adjustment. However, the results which indicate that distance estimation where photographs are stimuli is analogous to distance estimation on real terrain leads us to believe that performance which requires distance, speed, and time-of-arrival estimations can be trained on ATR with reasonable assurance of training transfer. If this proves to be the case, then ATR can become a useful tool in the early evaluation of a system design relative to predicting human performance on that system.

REFERENCES

Coynik, D. Movie making. Chicago: Loyola University Press, 1974.

Giannetti, L.D. Understanding movies. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

Gibson, E.J.; & Bergman, R. The effect of training on absolute estimation of distance over ground. Journal of Experimental Psychology, 1954, 48, 473-482.

Gibson, E.J., Bergman, R., & Purdy, J. The effect of prior training with a scale distance on absolute and relative judgments of distance over ground. Journal of Experimental Psychology, 1955, 50, 97-103.

Kraft, R.N., Patterson, J.F., & Mitchell, N.B. Psychological research on ATR: Formatting the visual material. The Army Research Institute: Alexandria, VA, in press.

Lippman, A. Movie maps: An application of the optical videodisc to computer graphics. Computer Graphics, 1980, 14, 32-42.

Teghtsoonian, R. & Teghtsoonian, M. Scaling apparent distance in a natural outdoor setting. Psychonomic Science, 1970, 21(4), 215-216.

Widen, H. Judgment of distance over ground in photographs: I. Institute of Military Psychology, (Swedish), B-83, 1973(a).

Widen, H. Judgment of distance over ground in photographs: II. Institute of Military Psychology, (Swedish), B-87, 1973(b).

AD-P003 360

STATIC vs. DYNAMIC PRESENTATION OF VISUAL CUES IN SIMULATED LOW LEVEL FLIGHT

Edward J. Rinalducci, Michael J. Patterson
School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332
and
Joseph DeMaio
Air Force Human Resources
Laboratory/Operations Training
Williams Air Force Base, AZ 85224

Abstract

The present study examined three visual display environments (i.e., a valley floor, a valley floor with walls, and a valley floor with walls and inverted pyramid terrain features) using different display presentation modes (i.e., slides, static video, and dynamic video). Both pilot and non-pilot subjects were employed. Differences between pilot and non-pilot sbjects were obtained for the accuracy of altitude estimation with the former being more accurate. Although the results were complex, both pilots and non-pilots, showed in general, an improvement in altitude estimation with the dynamic vs. the static mode of presentation and with increasing complexity of the visual scene. Resolution of the display image was also shown to be an important factor. The results of this study have relevance to the development of CIG and the evaluation of the simulator visual environment.

Introduction

Background

There has been a growing emphasis in pilot training on the simulation of low level flight. This has lead to an increased research effort into the nature of the visual cues required by pilots to maintain altitude. Pilots must be able to learn to judge their altitude with some accuracy in low

level flight operations.

The research described here examined the effects of a dynamic mode of presentation of a simulator visual environment compared to a static presentation on the ability of observers to accurately estimate altitude. Static scene presentation, as defined here, refers to the use of still photographs (35 mm slides) or video projection (i.e., without movement). Dynamic visual scene presentation refers to visual scenes with motion such as that presented through video projection (i.e., with movement). It has generally been assumed that the presentation of terrain features and depth cues in a dynamic scene is preferable to a static scene, especially when examining various aspects of visual space perception and visual information processing relevant to flight simulation. Empirical support on this issue, however, is lacking, and the present study provides relevant data. The objective of this research was to evaluate the nature of the mode of visual scene presentation and visual environment detail on the ability of observers to estimate altitude. It was hypothesized that altitude estimates in a dynamic scene mode of presentation would be more accurate than those made from a static scene presentation. Flight experience was varied by employing both pilot and non-pilot observers.

Methods

Materials

Stimulus materials consisted of 35 mm slides and video taped simulated flight segments. Slides were taken with a 90 degree field-of-view lens in the F-16 cockpit of the Advanced Simulator for Pilot Training (ASPT) located at Williams AFB, AZ. Video tapes of 5-second flight segments through three simulated visual environments were also made. Airspeed for the dynamic mode of presentation was 450 KIAS. Slides and tapes were made for eight altitudes between 50 and 400 feet and separated by equal log intervals (i.e., about 0.13 log units). There were three display modes consisting of slides, dynamic video, and static video (still frames from the dynamic portion of the flight segment). There were also three visual display environments or conditions. Display environment 1 was a textured valley floor, environment 2 was a valley floor with walls, and environment 3 was a valley floor with walls and inverted pyramids. The pyramids had black sides with white tops and were 50 feet tall with a mean distance between them equal to about 1500 feet.

Eight altitudes were presented for the three visual display environments for a total of 24 slides, static video frames, and dynamic video flight segments. The display environments were randomized within each display mode and presented three times.

Subjects

Twenty-one pilots (with approximately 400 to 3000 hours flying experience) undergoing simulator training at Williams AFB and 24 non-pilots (i.e., undergraduate students at Georgia Tech) served as subjects in this study.

Procedure

At the beginning of each session, the experimenter explained the purpose of the research, and that a sequence of 24 slides, static video frames, or dynamic video flight segments would be presented. Subjects were told that the presentatons would show three different simulator environments. Since none of the subjects had ASPT experience, they were told that the range of the altitudes would be 50 to 400 feet. Subjects were then given response sheets and told that when the first visual display enviroment was shown, they were to estimate the altitude above ground level (AGL). Estimates for subsequent visual environments were to be made relative to the first. That is, if the estimated altitude for the first visual display was 100 feet and the second visual display appeared to have been taken from an altitude twice as high, the second estimate should be 200 feet, and so on for succeeding visual displays. Thus, the psychophysical method employed was a variation of magnitude estimation using a free modulus technique (Engen, 1972). In a previous study (Rinalducci, DeMaio, Patterson, and Brooks, 1983) the method of magnitude estimation was shown to be a sensitive technique for the evaluation of simulator display systems. Each slide and each static video frame was presented for eight seconds. Each dynamic flight segment was presented for five seconds. Due to time limitations, the pilot subjects were only partially counterbalanced for the three display modes, while the non-pilot subjects were completely counterbalanced.

Results

The first display mode sequence was treated as practice and the altitude estimates from the second and third runs only were analyzed. A linear regression function was

determined relating log estimated altitude to log actual altitude for each display mode-display environment combination. The least squares technique was used to solve for the slope and y-intercept of the linear regression function. The dependent measure analyzed was the slope of the linear regression function which is the exponent of the power function obtained for each subject. The exponents were treated as individual data points. The values of the y-intercept were not analyzed. The data for both groups of subjects are shown in Table 1. In terms of the power function, an exponent (or slope of the log-log plot of the linear regression function) of 1.0 is indicative of accurate estimation of altitude. An exponent greater than 1.0 is indicative of expansion or overestimation of changes in altiude and an exponent of less than 1.0 is indicative of compression or underestimation of changes in altitude.

Table 1

Power Function Exponents of the Altitude Estimation Functions for the Static and Dynamic Visual Display Environments

Display Environment	Slides	Static Video	Dynamic Video
Valley Floor	0.20 (0.37)*	0.007 (0.29)	0.26 (0.72)
Valley Floor with Walls	0.27 (0.54)	0.17 (0.58)	0.38 (0.61)
Valley Floor with Walls and Inverted Pyramids	0.59 (0.78)	0.26 (0.55)	0.53 (0.84)

*Pilot data in parentheses

A 2 x 3 x 3 split-plot analysis of variance was performed on the data. The ANOVA for the data showed statistically significant differences (p less than 0.0001) for both the two within-subject variables (display mode with F = 40.63 and df = 2, 86 and display environment with F = 46.51 and df = 2, 86) and for the between-subject variable (flight experience with F = 31.76 and df = 1, 43).

Discussion

In general, pilot subjects were more accurate in their altitude estimates than non-pilot subjects, as is shown in

Table 1. It can also be seen that for pilots and particularly non-pilots, increasing the complexity of the visual display environment usually leads to an increase in the accuracy of altitude estimation independent of display mode. In addition, the dynamic mode of presentation appeared to result in more accurate estimates of altitude, especially when compared to the static video mode. The tendency for slide presentation to produce more accurate altitude estimation compared to static video suggests that clarity or resolution is an important factor in simulator visual displays.

Owen and Warren (1982) have indicated that individuals with no prior flight experience have still had exposure to global optical flow rates equivalent to those encountered in actual flight situations. However, for nearly all combinations of display modes and environments, the pilots of the present study did significantly better in altitude estimation than did non-pilots. One possible explanation may be related to the type of flow pattern experienced by pilots vs. non-pilots. There may exist some qualitative differences in the flow patterns experienced by both groups with pilots being sensitive to certain differences which result in increased accuracy. Therefore, the addition of optical flow in the dynamic mode may partially compensate for the lack of complexity and cues in the least complex display environment (valley floor alone). Pilots showed no significant differences between the least complex display environment and the most complex display environment (valley floor with walls and inverted pyramids) in the dynamic mode which was not the case for non-pilots.

References

Engen, T. Psychophysics II: Scaling methods. In J. W. Kling and L. A. Riggs (Eds.), Volume I: Sensation and perception, Woodworth & Schlosberg's experimental psychology. 3rd edition, New York: Holt, 1972, 47-86.

Owen, D. H., and Warren, R. Perceptually relevant metrics for the margin of aviation safety: a consideration of global optical flow and texture variables. Proccedings of the Conference on Vision as a Factor in Military Aircraft Mishaps, 1982.

Rinalducci, E. J., DeMaio, J., Patterson, M. J., and Brooks, R. Psychophysical assessment of simulator visual displays. Proceedings of the Second Annual Aviation Psychology Symposium, 1983, 489-494.

AD-P003 361

NOSC-Hawaii Perceptual Sciences Research Program

Ross L. Pepper
Naval Ocean Systems Center

Abstract

This paper describes NOSC-Hawaii's involvement in research in perceptual sciences, especially in teleoperator vision systems research and development. The major problem identified at the outset is that the rapid pace of technology has outdistanced the applied research community's progress in evaluating the utility of new techniques and devices designed to provide sensory information to the human operator.

The approach taken by NOSC-Hawaii is to develop a methodology and supporting test facilities which will facilitate the development of a general theoretical model capable of describing and accurately predicting all aspects of teleoperator system performance across a wide variety of operating conditions and tasks.

To date, NOSC-Hawaii has developed a capability to conduct human performance evaluations of visual system components that exist in no other location. Our approach, which employs a method of analysis which we have termed the "display-performance transform," is unique and comprehensive. Its simplicity enables clarification of interacting system elements, provides standards of evaluation, and promotes conceptual dialogue on theoretical issues.

A description of past work conducted by the Hawaii Laboratory is presented. The value of this section is that it establishes for the reader the continuity of NOSC's efforts and provides ready access to previously published work.

The final section of this paper addresses the need for continued cross disciplinary research and development. In it, the directions taken by NOSC-Hawaii are revealed. The role of the Perceptual Sciences Group is clearly underlined and the future potential for expanding man's performance using teleoperator systems is described.

PROBLEM: Advances in teleoperator technology have outpaced our research efforts to identify and assess the usefulness of new products and procedures available to enhance performance.

Technology Growth

There are several issues here that require clarification. The first is the technology problem. The creative productivity of the 70's and 80's has resulted in the emergence of a robotics and teleoperator technology that involves many applied sciences (there is no common terminology, standard testing procedures, etc.). Furthermore, most of the technology has been designed to satisfy other needs. The roots of this emerging technology are as disparate as one can imagine: From man's vision of space exploration, and subsequent need for unmanned space probes, the first truly autonomous vehicles were developed. Earlier but now parallel development has occurred in the Navy's efforts to work in the deep ocean environment. These two environments represent the extremes in hostile working conditions, and both have spawned a variety of autonomous and teleoperated vehicles to satisfy mission-specific objectives. The problems and applications

have not diminished. Today many scientists and engineers sense a new vitality, a new interest to be served in these traditional arenas.

Task Requirements Problem

A second issue, which goes hand-in-hand with the technology growth problem, is of special significance to those involved in the design and development of teleoperator systems. This issue has to do with the task requirements problem.

In order to evolve teleoperator systems which can perform increasingly complex tasks, requiring real-time intelligent analysis of a variety of sense data, we must understand the dynamics of the information-gathering, processing, decision, and reaction processes in much greater detail.

We must understand the opportunities and complexities involved in augmenting man's performance, in sensing information, as well as in responding to various task contingencies. We must evaluate man's continuously evolving role in relation to the increasing capabilities provided by advances in computer and cognitive sciences. And, at all times, we must strive to design teleoperator systems to fit the human operator by employing analytic approaches which are teleoperator friendly.

OBJECTIVE: The broad objective of our group is to develop an understanding of the dynamic interaction which occurs between the human operator and the sensors, controls, and displays of remotely manned systems (teleoperators). The emphasis is placed on establishing a fundamental understanding of man's perception of information received from sensors and displays and his performance using control systems and their associated manipulator devices, in order to produce increasingly intelligent, skilled performance across a variety of specified tasks. In order to accomplish this objective, we have begun to develop a general model of remote operator performance that will ultimately have broad predictive ability. A fundamental prerequisite for the development of this model is a data base of performance derived both from empirical tests as well as from predictions based on existing models of perception and motor-skill performance. Such models are generally described in the psychological literature (see Adams, 1967; Bilodeau, 1966).

The scope of our group's interests necessarily includes elements of basic physiology, sensory psychology, and cognitive psychology as well as some of the more traditional engineering approaches in control theory and mechanical systems.

Background

NOSC has a long history of involvement in research to support the development of remotely manned work systems and undersea vehicles. Initial work in visual systems research was funded by the Office of Naval Research. These efforts were directed toward conducting an evaluation of the utility of stereoscopic TV compared to conventional TV as an aid to operator performance.

Until recent years, it was widely held that stereoscopic, as compared to monoscopic, viewing systems contributed little to the successful performance of remote manipulation tasks (Pepper and Cole, 1978; Kama and DuMars, 1964). This general conclusion stood in contradiction to a large body of psychophysical evidence showing that object location, form detection, and recognition performance were greatly increased with binocular compared to monocular viewing conditions (Graham, 1965). With this contradiction in mind, we performed a detailed analysis of display system variables affecting operator performance in undersea teleoperators and concluded that performance is determined by a complex interaction of numerous factors including the spatial information available at the

remote site, its transmission and display to the operator, manipulator capability, task demands, and human operator capabilities including perceptual and perceptual-motor learning skills. Our experimental efforts to date have focused on the effects of visual display parameters on overall system performance.

Our earliest experiments (Smith, Cole, Merrit, and Pepper, 1979) assessed the effects of task factors, learning, and visibility conditions at the remote work site on purely perceptual as well as perceptual-motor (manipulative) performance. Using a simple target positioning task with highly practiced subjects to minimize learning effects, mono and stereo TV performances were measured at three levels of visibility. As predicted, stereo performance was superior to mono under all conditions tested. Performance using both mono and stereo displays was affected by degraded visibility. In another experiment, a messenger-line feeding task was employed which required subjects to thread a length of rope through a complex set of hoops. Results of this experiment indicate that the stereo advantage increased with decreases in visibility. A similar pattern of results was observed when performance time was taken as the dependent measure. Thus, performance of a complex task was both faster and more accurate under stereo viewing conditions, with stereo providing greater advantage under degraded viewing conditions.

Among our most recent studies are those investigating the possible advantages of using head movement in conjunction with bench-mounted stereo displays in order to enhance stereoacuity (Cole, Pepper, and Pinz, 1981). In these experiments, stereoacuity was measured with a modified Howard-Dolman apparatus which had different-sized standard and comparison rods. Since our TV cameras remained in a fixed position throughout testing, observer's lateral head movements relative to the stereo display did not produce true motion parallax cues to depth. However, the apparent motion of objects in the scene is, like true motion parallax, proportional to their distance from the convergence plane of the remote cameras but in the opposite direction of that encountered with true motion parallax. Despite the illusory nature of this effect, it seemed reasonable to suppose that the relationship between the apparent distance of cameras and the degree of "pseudo-motion parallax" of those objects could be evaluated by the visual system in much the same way that true motion parallax is evaluated. Our original study of this phenomenon, however, did not confirm this hypothesis. While stereoacuity under stereoscopic viewing conditions was superior to that under monoscopic viewing conditions, pseudo-motion parallax cues provided by operator head movements were found to neither enhance nor degrade stereoacuity above the level of performance associated with the use of stereo viewing conditions alone.

Another area of interest in our recent studies (Pepper, Cole, and Spain, 1983) is lateral camera separation, which results in a magnification of the retinal disparity cue to depth in stereo displays. With judgments of the relative depths of luminous disks set against a dark background, we have found that wide camera separations can actually enhance stereoacuity beyond levels found under comparable direct viewing conditions. Whether this advantage will persist for more complex visual or manipulative tasks will be assessed in future studies using more naturalistic scenes.

In our most recent studies (Pepper, Cole, Spain, and Sigurdson, 1983) we have employed a helmet-mounted stereo TV display (HMSD) and an isomorphic head-motion tracking system developed by NOSC engineers which allows us to convey true head motion parallax cues to an operator. Our preliminary results with this system have been encouraging. Stereoacuity for both simple and complex perceptual tasks is consistently superior with head-motion camera-coupling under both

stereoscopic and monoscopic viewing conditions. However, the complexities of the particular head-motion camera-coupling system which we are using at present are considerable. Our HMSD might also place additional demands on the operator which may or may not be offset by performance gains associated with the added degree of complexity and sophistication.

While these results have had significant application to the development of remote viewing systems for specific applications, they have exposed two glaring deficiencies in our understanding of such systems. The first is a lack of a complete, theoretical model of human spatial vision. Some theories have been proposed to deal with limited aspects of visual perception (see Gogel, 1977), but to date no one formulation is comprehensive enough to accommodate all of the empirical evidence. An even greater deficiency results from the discontinuity between direct and remote viewing. Under direct vision, two parameters, convergence angle and distance to a reference object, determine retinal disparities. In the stereo TV display condition, retinal disparities are determined by seven parameters: occular convergence angle, eye to display screen distance, camera to reference object distance, camera separation, camera convergence angle, camera field of view (i.e., magnification), and display screen width. Viewing a scene directly and viewing it through a display introduces a whole new level of factors that can affect visual performance. One serious consequence of this inadequacy is the fact that existing psychophysical data do not provide accurate predictions of visual performance employing remote displays, and therefore are not an adequate basis for making informed engineering decisions about features to be incorporated into teleoperator displays.

In the absence of any great progress in the development of a psychophysics of teleoperator systems, the current advances in electronic technology have produced increasingly sophisticated devices which can be employed to extend man's performance using remotely operated systems. Additional advances in computer sciences and cognitive psychology have also produced a need to understand the relationship (both perceptual and cognitive) which exists and/or can be developed to augment man's performance with teleoperator systems. In order to attain the goal of a detailed specification of teleoperator system components and parameters, we believe that it is necessary to assess the effects of each of the factors of the man-machine system both independently and in various combinations.

Approach

In the preceding section, I have tried to identify some of the issues which need to be addressed in order to make progress toward the goal of a mature teleoperator technology. One of the major milestones which will enable us to attain that goal is the development of a model of teleoperator performance. The utility of a good model cannot be overstressed; it can clarify ambiguity and lead to predictions which might never be entertained.

The first step in model development is the necessity to obtain a clear, ordered set of data. We have developed a program of research which includes a laboratory with the capability for remote task simulation as well as for basic research in the visual perception of objects in remote environments. We have evolved a method of analysis in order to understand how man's perception of displayed information from the remote world is altered, compared to real world information. We have termed this method the display-performance transform. The basic tenent of this method is that the relationship between displayed information and man's performance is lawful and ultimately predictable. To date, however, the level of understanding of these transformations is currently imprecise. Often, the performance transformation which occurs can be traced to

specific camera or display qualities. However, due to the complexity and interactive nature of the hardware components available to produce information to the operator, as well as the wide range of perceptual motor-skill performance required of the operator, the ability to predict this transformation is limited.

In summary, the display-performance transfer method is an empirical approach which produces standard measures of performance. These measures are the fundamental data necessary to reveal how various system elements interact. They additionally provide a basis for the development of a general model of teleoperator systems. Ultimately, the model development will enable tests of conflicting theoretical or conceptual interpretations of man's remote performance. It will also provide a practical, predictive basis for the selection of newly developed hardware devices and/or teleoperator features. At yet another level, this approach also produces an objective basis for determining costs versus benefits associated with teleoperator system components.

To my knowledge, no other laboratory has attempted to develop a systematic, multi-level test and evaluation facility. Ours has been limited to visual perception. We intend, over the next several years, to broaden our efforts to include manipulator and control systems so that the entire remote system is incorporated within our analytic framework.

References

Adams, J.A. Human Memory. New York: McGraw-Hill, Inc., 1967.

Bilodeau, E.A. (Ed.). Acquistion of Skill. New York: Academic Press, 1966.

Cole, R.E., Pepper, R.L., and Pinz, B.L. The Influence of Head Movement Parallax on Perceptual Performance Under Direct- and TV-Displayed Conditions. NOSC TR 678, May 1981.

Gogel, W. The Metrics of Visual Space. In Epstein, W. (Ed.). Stability and Constancy in Visual Perception: Mechanisms and Processes. New York: John Wiley and Sons, 1977, 504-547.

Graham, C.H. (Ed.). Vision and Visual Perception. New York: John Wiley and Sons, 1965.

Kama, W.N. and DuMars, R.C. Remote Viewing: A Comparison of Direct Viewing, 2-D and 3-D Television. Report AMRL-TDR-64-15, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, OH, February 1964.

Pepper, R.L. and Cole, R.E. Display System Variables Affecting Operational Performance in Undersea Vehicles and Work Stations. NOSC TR 269, June 1978.

Pepper, R.L., Cole, R.E., and Spain E.H. The Influence of Camera Separation and Head Movement on Perceptual Performance Under Direct and TV-Displayed Conditions. Society for Information Display, International Symposium, San Diego, CA, May 1982.

Pepper, R.L., Cole, R.E., Spain, E.H., and Sigurdson, J.E. Research Issues Involved in Applying Stereoscopic Television to Remotely Operated Vehicles. Proceedings of the International SPIE, Vol. 402, Geneva, Switzerland, April 1983, 170-174.

Smith, D.C., Cole, R.E., Merritt, J.O., and Pepper, R.L. Remote Operator Performance Comparing Mono and Stereo TV Displays: The Effects of Visibility, Learning, and Task Factors. NOSC TR 380, February 1979.

AD-P003 362

Factors Affecting Visual Performance with Stereoscopic Television Displays

by

Edward H. Spain
Naval Ocean Systems Center

ABSTRACT

A geometrical model describing retinal disparities produced by stereoscopic (stereo) TV systems is experimentally tested for various configurations of camera separation and magnification. Operator visual fatigue is also measured. Although varying configurations affect perceptual performance, they do not affect it in a manner consistent with the simple geometrical model. Observers appear capable of perceptual adaptation, benefitting from enhanced disparities when they are available. No differences in visual fatigue were found for any of the camera configurations tested.

INTRODUCTION

Stereopsis involves simultaneously viewing some aspect of the external world from two slightly separated vantage points and perceptually blending these two distinct perspectives into a unitary mental representation of external space. Though not a prerequisite to adequate spatial perception in many real-world situations, stereopsis is frequently a powerful adjunct to other visual spatial cues, particularly with regard to precise localization of objects in depth. Because of this stereo TV is an important feature of teleoperated systems which must interact efficiently and intelligently with their environments. While it is theoretically possible to exactly reproduce the pattern of stimulation available at a remote site, technical limitations of video systems have thus far fallen short of this ideal. Efforts are underway to improve video hardware as well as to systematically vary the perceptual cues in video imagery of remote environments in order to enhance an observer's perception and ultimately the operational efficiency of remote teleoperator systems. Given worldwide interest in stereoscopy for well over a century, it is surprising that the available scientific/technical literature of visual performance with stereo TV displays is so limited. Conclusions and technical guidelines in many of the available reports are based on faulty methodologies as well as untested and questionable assumptions regarding the perception of spatial relationships in remote scenes. Perceptual testing of sophisticated video systems often requires a concerted effort on the part of display engineers and perceptual scientists and is usually both time consuming and expensive. Many designers of stereo TV systems have relied solely on their own immediate subjective impressions as the sole means of assessment. This

overly simplistic approach often confuses functional and aesthetic aspects of image quality, fails to provide a basis for generalization of findings from the laboratory to the operational environment, and provides no useful information regarding the nature of the perceptual processes involved. Scientifically controlled investigations of space perception with stereo displays are few and of recent origin.

Under direct viewing conditions, unfusable images for central vision are almost never encountered because observers automatically adjust the optical axes of their eyes to produce fused images. A typical observer with no ocular abnormalities can readily adjust his eyes to fuse objects from 6 inches out to infinity. With stereo TV systems, there is frequently a danger of providing the operator with disparities that exceed the limit for fusion and stereopsis. Unlike the human eyes, which are relatively fixed in the head with respect to directly viewed objects, stereo camera systems are variable in terms of camera separation and image magnification. Like the eyes, these systems can be converged and diverged, but the range of values within which they can be adjusted is obviously much greater than that of the eyes. Thus, there are many more possible combinations of retinal disparities, object sizes, and textural gradients under stereo TV viewing conditions than are possible under direct viewing conditions. Stereopsis with stereo TV systems is similar to stereopsis under direct viewing conditions insofar as retinal disparities are a potent cue to depth perception, but there are several important differences between direct and TV viewing conditions. These differences can exert strong effects on the perception of depth with video displays and involve the number of geometrical parameters determining disparities, information carrying capacities of present-day video systems, and perceptual cue conflicts.

Two parameters, convergence angle and distance to a reference object, determine retinal disparities under direct viewing conditions. In the stereo TV display viewing situation, retinal disparitiess are determined by 5 additional parameters. Precise geometrical formulas relating these parameters to retinal disparities are well-known (see Shields, Kirkpatrick, Malone, and Huggins, 1975). Taken together, they provide a precise mathematical model for estimating retinal disparities from objective linear and angular measurements of the observer, the viewing station, the TV hardware configuration, and distances between sensors and objects in the remote environment. Of particular interest are the parameters of camera magnification and camera separation, both of which can be readily manipulated in stereo systems in order to increase or decrease image sizes and retinal disparities. As camera magnification increases, the angular sizes of displayed objects and their associated disparities increase. Increasing camera separation magnifies disparities associated with objects without altering their sizes, and also brings about an enhancement of perceived depth.

The fundamental flaw of all purely geometrical formulations of space perception with stereo TV displays is an overemphasis on retinal disparities as the dominant cue to accurate depth perception and a tacit disregard for other powerful spatial cues. Depending on the combination and distribution of these cues within a remote scene, they may reinforce, mitigate, or override perceptual information provided by retinal disparities alone. Stereo TV viewing conditions constitute a special case of space perception, a case in which the pool of cues may be substantially different from those available under direct viewing conditions in physically subtle but perceptually powerful ways. Because of this one cannot assume that findings reported in the traditional psychophysical literature are applicable to stereo TV viewing. Controlled empirical investigations of the type described below are essential to the advancement of TV stereoscopy. Since the geometrical model describing retinal disparities frequently plays a prominent role in the design of present-day stereo TV systems, it is important to test the model for its ability to predict perceptual performance. In particular, how do variations in camera magnification and camera separation (both of which exert highly predictable effects on retinal disparities) independently and interactively influence the perception of depth intervals in a remotely televised scene? Do variations in these parameters have differential effects on visual fatigue? The following experiment was conducted to provide preliminary answers.

METHODS

Four observers, two females (ages 33 and 22) and two males (ages 33 and 18), participated in an experiment designed to study the main and interactive effects of camera separation and lens magnification on the perception of depth intervals in a televised space. All were screened for ocular anomalies. Prior to testing, each observer was given no less than five hours of practice in making depth judgments with stereo TV displays, though two observers had approximately ten times as much practice. Testing consisted of thirteen one-hour sessions with each session measuring observer performance under a single set of TV system parameters derived from a factorial crossing of a) four levels of camera separation (0 in. (monoscopic), 1.25 in., 2.5 in., 7.5 in.), b) three levels of magnification (1X, 2X, and 3X), and c) six objecti.e depth intervals (0, 2, 4, 6, 8, 10 in.).

Within a single testing session three brief measures of visual efficiency were administered immediately before and after testing with the stereo TV system. Observers first answered a short visual fatigue questionnaire. Next, they were administered a timed near-far test of visual acuity which required rapid alternations of convergence and accommodation of the eyes between a Landolt test object .5 meter distant and a target 6 meters distant. Gaps in both near and far test objects subtended a visual angle of 1.5 arcminutes at the viewing station. Next,

observers adjusted the flicker rate of a circular patch of red light set within a darkened background. On four trials they started at a clearly visible flicker rate of 25 Hz and adjusted frequency to the critical flicker fusion frequency (CFF). CFF was measured because many studies have suggested its usefulness as an indicator of central nervous system fatigue.

Visual performance was next measured across 60 experimental trials with a polarizer stereo TV display similar to that used in our previous studies (Cole, Pepper, and Pinz, 1981). The televised remote scene consisted of two black rods of equal diameter which were positioned in front of an illuminated but patternless white background. Distance from the cameras to the null point around which the rods were adjusted in depth was 2 meters. The rods were laterally separated by 12.7 cm. For each trial, observers first verbally reported which of the two rods (left or right) appeared closer. Next, they reported how many inches the rods appeared to be separated in depth. Finally, they adjusted the depth of a sliding peg relative to a fixed peg so that it matched the apparent depth between the rods in the televised scene. The latter measure is referred to as a haptic adjustment. Observers were provided immediate verbal feedback as to how far and in what direction (i.e., "short" or "long") their haptic adjustment errors were made. Immediately following stereo TV testing, the observer was again administered the CFF measure, the near-far acuity test, and the visual fatigue questionnaire.

RESULTS AND DISCUSSION

Testing yielded 5 sets of scores for preliminary statistical analysis: a) error scores for verbal reports of perceived depth, b) error scores for haptic adjustments indicating preceived depth, c) scores for the visual fatigue questionnaire, d) response times for the near-far acuity test, and e) fusion thresholds for the CFF measure. All analyses were repeated-measures analyses of covariance. Assumptions of the analyses were empirically tested and adjustments were made to F-ratio tail probabilities when appropriate. Since similar patterns of results were found for haptic adjustments and verbal reports of perceived depth, only results for haptic adjustments are reported here. Geometrical theory predicts that as values of camera separation and magnification increase, disparities increase and therefore the perceived depth associated with fixed depth intervals should also increase. The analysis revealed significant main effects for camera separation ($F=73.13$, $p<.001$) and the interaction of camera separation with lens magnification ($F=8.43$, $p<.02$). Camera lens magnification did not exert a significant main effect on perceptual performance even though, according to the geometrical theory, it did have a multiplicative effect on disparities in the imagery. The analysis of visual fatigue primarily concerned with pre-post comparisons of questionnaire reports of eyestrain, the near-far acuity test, and CFF measures failed to show any statistically significant

differences.

Camera separation exerted an influence on both verbal and haptic measures of perceived depth but did not do so in a manner consistent with the simple geometrical model of retinal disparities. When disparities increased beyond "normal" levels as a result of increasing camera separation, perception of depth in the remote environment became more accurate, not more distorted. Rather than responding in a simple, stimulus-bound manner, observers appear to have actively adapted to stereo viewing conditions on the basis of other information available within the testing situation. That is, foreknowledge of the range of possible depth intervals in the scene and immediate feedback following each trial may have been used to gauge the relationship between disparities in the stereo images and depths in the televised scene.

Magnification was not found to have a significant influence on perceived depth. Although it did affect disparities, it did so to a lesser extent than camera separation. It did not disrupt the effects of camera separation. Targets used in this experiment were unpatterned, and so a major benefit of magnification, that of increased detail resolution, was not a significant factor in determining depth perception.

Measures of visual efficiency taken before and after testing with the stereo TV system failed to show any statistically significant differences in visual fatigue across the set of values for camera separation and magnification tested. It should be noted, however, that great care was taken in balancing and aligning the two channels of the TV system and that average testing time on the stereo display was only about 20 minutes. Thus, moderate exposure to "unnatural" disparities produced by a carefully tuned stereo TV system do not appear to engender very much if any visual fatigue in observers. Longer exposure times will be necessary to reveal differential effects of stereo viewing system parameters on visual fatigue. How these factors influence the perception of depth when scenes of greater complexity are viewed through a stereo TV system is a topic currently under investigation in our laboratory.

REFERENCES

Cole, R.E., Pepper, R.L., and Pinz, B.L. The influence of head movement parallax on perceptual performance under direct and TV-displayed conditions, San Diego: Naval Ocean Systems Center TR 678, May 1981.

Shields, N., Kirkpatrick, M., Malone, T.B., and Huggins, C. Design parameters for a stereoptic television system based on direct depth perception cues. Proceedings of the Human Factors Society 19th Annual Meeting, Oct 1975, pp 423-427.

Interactions of Motion Perception with Pattern Perception

by Captain William Marshak, Department of
Behavioral Sciences and Leadership
U.S. Air Force Academy, Co

Abstract

Perception of motion and spatial pattern are closely intertwined. The reaction time for seeing changes in direction are longer for certain spatial arrangements of moving dot patterns than for others. When a stationary pattern of dots overlaps a moving pattern of dots, the amount of adaptation measured by motion aftereffect is significantly attenuated. This is particularly true when collisions occur between dots. An explanation of these motion - pattern interaction is offered in terms of a vector model of motion perception.

Introduction

The visual environment both inside and outside the aircraft cockpit is becoming increasingly complex, especially in the domain of motion perception. Development and testing of a theory of motion perception which can handle this complexity and allow prediction of potentially disastrous misperceptions is increasingly urgent. Earlier work (Marshak, 1981) proposed a theory for motion perception called the vector model, which is a form of multichannel motion processing loosely based on known physiology. Unlike some perceptual constructs, the vector model is specific how the multiple channels will interact and makes unambiguous predictions about the perceptual process.

Three stages of motion information processing are proposed in the vector model. The first stage consists of motion sensitive units which are maximumly active with motion in one direction, but respond less to adjacent directions of motion. The activity of these units are dependent on the temporal frequencies generated by motion (Pantle, 1974) which are the combined result of the spatial frequency or pattern of the moving target, its direction of motion and the velocity of motion. The temporal frequencies along orientations adjacent to the direction of motion moderate with a cosine function.

These motion sensitive units are not independent. There is lateral inhibition among the units. Units increasingly inhibit adjacent units with increasing angular difference up to approximately 20 degrees and then the amount of inhibition falls off. Such a mechanism is a form of neural sharpening which enhances the accuracy of perceived direction. This mechanism is revealed by the repulsion of different directions of motion viewed overlapping observed by Marshak and Sekuler (1979) and Mather and Moulden (1980).

Finally, the vector lengths greater than threshold levels are combined in a vector summation process to determine the perceived direction of motion. The second stage neural inhibition acts as a limiter of the range over which summation occurs, allowing distinguishing of multiple directions.

Consider two possible consequences of the vector model. The stage one generation of activity in motion sensitive units is dependent in part on the pattern of the moving object, more specifically the spatial frequency contribution to the distribution of temporal frequencies. If spatial energy were to be concentrated along a particular orientation, the temporal frequencies generated along that orientation may be greater than along the adjacent directions of motion. This would result in a measurable increase in reaction times to changes in direction off the energy loaded orientation.

A stationary pattern affects the motion sensitive elements of the vector model in the following fashion. Some activation in all units results from stationary stimuli. summation of this activity in all directions would be a zero sum or perceived stationarity. There would also be inhibition generated among all the units as well, but it to would balance out. If a stationary pattern were overlapped with a moving pattern, the global inhibition might depress the activation of the units responding to the motion. Long duration motion combined with a stationary pattern would result in less adaptation than without the stationary pattern. The result on motion aftereffect would be an attenuation of the aftereffect.

Experiment 1

Each trial consisted of a stimulus pattern moving in a direction of 180 degrees (polar) or leftward for a randomly chosen interval between 1.5-2.5 seconds at four degrees of visual angle per second and with a suprathreshold contrast. After this 180 degree motion, the whole pattern abruptly changed direction to a new one. This new direction was plus or minus 10,20,30 or 40 degrees from 180. One independent measure was the size of directional change.

The second independent measure was the kind of moving target. Three different stimuli were employed. The first was a random dot pattern containing 600 points. Spatial frequency energy is equally distributed in random dot patterns. The second pattern were dots arranged in parallel vertical lines. Inter-dot distance was sufficient that the individual dots were discernable at the viewing distance. Each dot subtended .028 degrees of visual angle and the distance between dots were .070 degrees. These vertical lines of dots appeared at regular horizontal intervals of .383 degrees. The lines extended vertically from a random starting position on the Y axis. These vertical lines were one-third the screen's length so that the ends of the lines were visible, although they may wrap around the vertical axis. Spatial energy in this stimulus should be loaded perpendicular to the verticals, or along the axis of initial motion. The third stimuli were lines constructed similarly to the vertical lines. Their start points were distributed equal distant along the horizontal axis of the screen. Unlike the 'vertical line' stimuli, the orientation of the lines are random and appeared as random line or 'pick-up stick' stimuli.

Four observers served in both the experiments. Three were naive and inexperienced in psychophysical experiments. The fourth observer was myself. Vision of the observers was 6/6 or corrected to that acuity. Each observer saw each of three stimuli move in the eight directions for four repetitions. Thus, the experiment consisted of 96 trials.

Results

The mean reaction time for all four subjects for each stimulus and at each direction of motion is plotted in figure 1. The data was analyzed using a repeated measures analysis of variance.

The results are quite striking. There is a significant direction effect ($F=12.936$, $df=7/21$, $p<.001$) . It takes longer to see a small direction change than it takes to see larger changes. Reaction time is longer with 10 degree changes in direction, but seems to asymptote by the twenty degree direction change.

The is also a significant stimulus effect ($F=71.548$, $df=2/6$, $p<.001$). Reaction times to random dots and non-parallel line patterns were virtually the same . Each exhibits the elevation of in reaction time with 10 degree changes. The whole reaction time curve for

parallel vertical lines was elevated across all but the most extreme (40 degree) direction change. This is consistent with the vector model of motion perception proposed in my dissertation and the research proposal. The spatial energy of the parallel vertical lines is loaded along the horizontal orientation, coincident with the direction of motion. When the stimulus moves in the new direction of motion, the spatial energy distribution produces significant temporal frequency energy along in the leftward direction masking the change of direction and elevating the reaction time.

There is also an interaction between stimulus and direction ($F=4.435$, $df=14/63$, $p<.001$). The range of directions of spatial influence on reaction time is much larger for the parallel vertical line stimulus than for the other two stimuli. The effect tapers off gradually with increasing directional change. Temporal frequencies generated by motion generate frequencies in directions adjacent to the direction of motion. The frequencies fall off with increasing difference or direction by a cosine function. The gradual tapering in elevation of reaction time with increasing directional change observed with vertical line stimuli is consistent with the cosine function of temporal frequency attenuation.

Experiment 2

The second experiment uses a different method to reveal spatial influences on motion perception. Instead of using motion perception, the experiment studies the closely related phenomenon of motion aftereffect (MAE). At issue is whether stationary patterns can influence the amount of MAE experienced.

The independent variable in this experiment was the adapting stimuli. Three moving stimuli were used to adapt four observers. The first stimuli consisted of 300 moving random dots. The second adaptation stimulus consisted of moving 300 random dots superimposed on a stationary pattern of 300 pseudo-random dots. The stationary dots were positioned along the horizontal axis so that collisions with the moving pattern occurred, but not simultaneously. Thus, each moving dot collided with one stationary dot once during each screen transit. The third adapting stimulus consisted of moving 300 random dots superimposed on another stationary pattern of 300 pseudo-random dots. This stationary dots never collided with the the horizontally moving dots.

A control independent variable was direction of motion. Motion leftward and rightward were both employed to minimize adaptation carrying over trials . The order of stimuli and directions was random but equal in occurance.

Size, contrast and velocity of the dots were the same as the first experiment. Adaptation consisted of one of the three stimuli moving at 4 degrees per second for 25 seconds. The same four observers were instructed to maintain their gaze on the centrally located fixation point during adaptation and resist the powerful tendency to follow the moving pattern. This was immediately followed by presentation of the dot pattern that had moved, now stationary, for 30 seconds. The observers watched the MAE experienced with the now stationary dots and recorded the end of MAE with a button press. They continued to watch the pattern the full 30 seconds to dissipate any residual MAE. A five second inter-trial interval separated each trial, during which the screen contained only the central fixation point. Each subject made 18 observations using the three stimuli at each of two directions with two repetitions.

Results

The resulting durations of MAE for the three stimuli are plotted in Figure 2. The

direction and repetition effects were not significant, as expected. The main effect for stimulus indicates that the presence of a stationary coincident pattern significantly reduces the amount of MAE experienced by the observers ($F=48.979$, $df=2/8$, $p<.001$). This effect is largely a result of stationary dots reducing the amount of adaptation that took place. The overlapping dots produced the least adaptation as expected, but the difference between this condition and non-overlapping dots was small. This reduction in adaptation is caused by stationary dot patterns exerting a powerful inhibitory effect on the motion sensitive elements resulting in less adaptation.

Discussion

The vector model is capable of accounting for two very different findings in the domains of immediate perception and aftereffects. Experiment 1 indicates that powerful loading of spatial energy in the motion of direction of a moving target can result in increased latency in seeing direction changes. Experiment 2 demonstrates how the presence of stationary patterns during adaptation can attenuate motion aftereffect. Exact modeling of the excitatory, inhibitory and summation mechanisms will allow precise prediction of the accuracy of motion perception in the laboratory and the environment.

The increased latency to changes in direction of motion supports the vector model, but is likely only a laboratory phenomenon. It takes a substantial amount of spatial energy to create the reaction time increase. Such dramatically shaped objects are rare in the environment. Since most direction perception in the aviation environment probably are the result of judging displacement of the target, it is unlikely that the low level processes proposed in the vector model are important.

The attenuation of motion aftereffect does explain why in our dynamically moving world, we do not experience more motion aftereffect. Pilots who fly at low level in bubble canopies which lack stationary elements in their visual environment may be susceptible to MAE. They may experience severe motion aftereffect which the present results suggest can be attenuated by etching stationary grid patterns inside the canopy.

References

Marshak, W. and Sekular, R. Mutual repulsion between moving visual targets, Science, 1979, 205, 1399-1402.

Marshak, W. Perceptual integration and differentiation of directions of moving patterns, unpublished dissertation, Northwestern University, 1981.

Mather, G. and Moulden, B. A simultaneous shift in apparent direction: Further evidence for a "distribution-shift" model of direction coding, Quarterly Review of Experimental Psychology, 1980, 32, 325-333.

Pantle, A. Motion aftereffect magnitude as a measure of the spatio-temporal response properties of direction-sensitive analyzers, Vision Research, 1974, 14, 1229-1236.

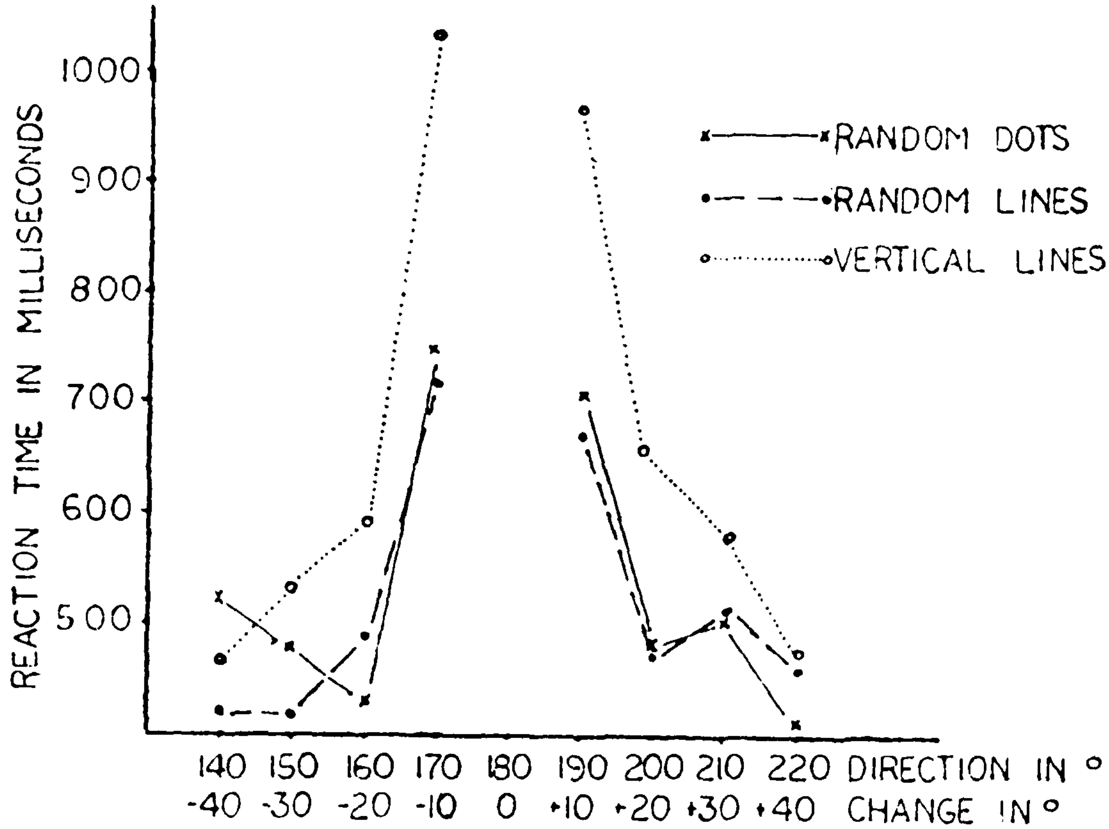


FIGURE 1.

REACTION TIME TO CHANGES IN DIRECTION

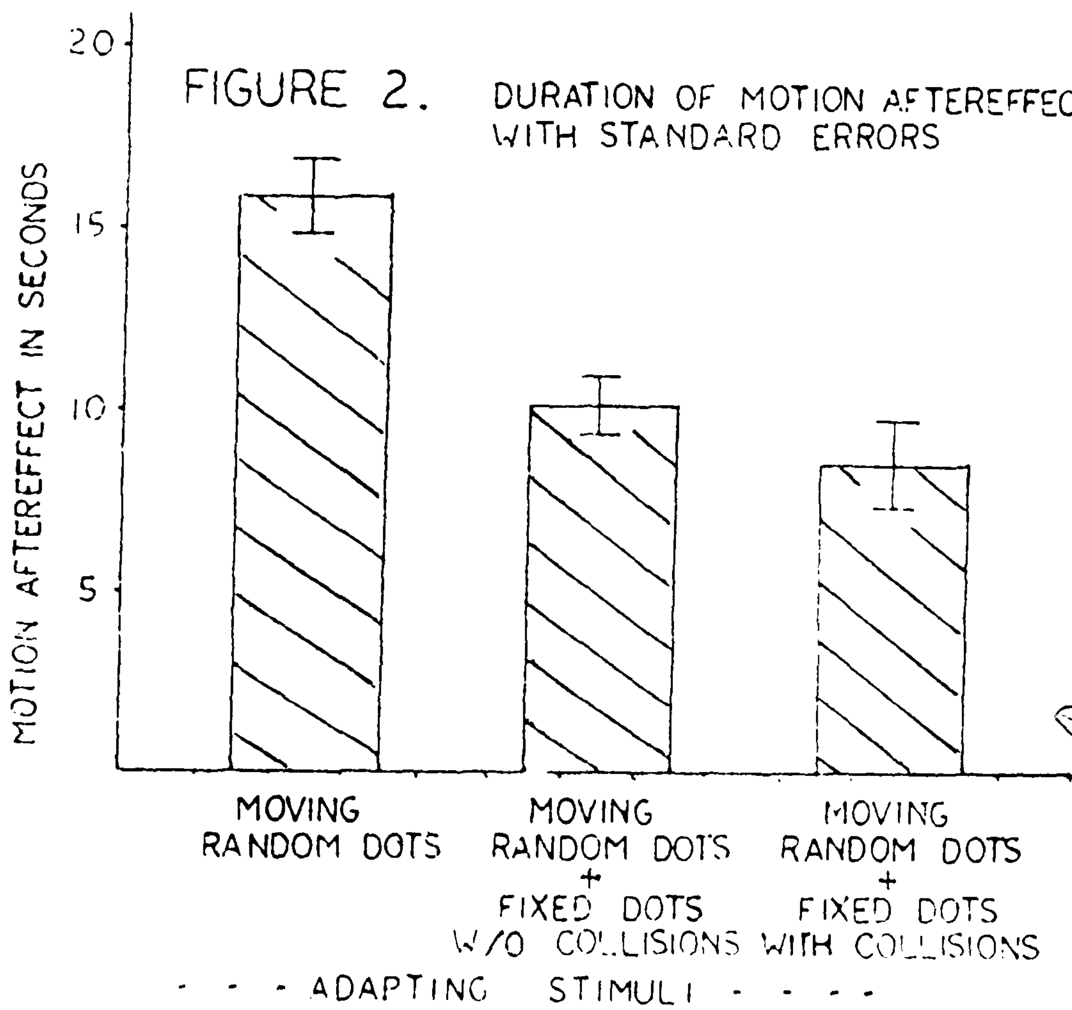


FIGURE 2. DURATION OF MOTION AFTEREFFECT WITH STANDARD ERRORS

PANEL SESSION

Cognitive Strategies Training

Larry W. Brooks, Chairperson
U.S. Army Research Institute for the
Behavioral and Social Sciences

Abstract

The purpose of this panel discussion is to present current DoD in-house research on cognitive strategies training. Specifically, the first presentation will be an overview of learning strategies training within a military setting along with examples of applied research in this area. The next three presentations will cover three areas of basic research in this domain: direct comprehension strategies, direct utilization strategies, and indirect affective strategies. The four presentations will be critiqued by a discussant.

Introduction

Military personnel are required to process large amounts of complex information in order to adequately perform their jobs. Because of the demands placed on the cognitive skills and abilities of military personnel, military research psychologists have initiated a growing number of DoD in-house efforts to investigate individuals' acquisition and utilization of cognitive strategies.

Cognitive strategies can be referred to as those procedures and techniques which allow a learner to gain and use new information in an effective manner. Strategies may either have a direct influence on how a learner acquires and uses information, or they may have an indirect influence on the learner's manipulation of existing information.

Direct cognitive strategies may be used primarily for the acquisition and comprehension of new information, or for the utilization of already acquired information to perform a given task. Comprehension and acquisition strategies would include techniques such as outlining text for studying or the creation of mnemonic elaborators to remember a list of items. Utilization strategies would involve the use of problem solving techniques and the use of strategies to facilitate the transfer of knowledge from one domain to another. Indirect strategies typically concern the affective and motivational state of the learner. These strategies would include skills for managing anxiety and techniques to increase positive feelings toward a particular learning situation.

The purpose of this panel discussion is to present current DoD in-house research in each of these three areas of cognitive strategy training: acquisition and comprehension strategies, utilization strategies, and indirect strategies. The first presentation by Ruth Phelps serves as an overview to cognitive strategies training in the military, and introduces two areas of applied research in this area. In

the second presentation Larry Brooks discusses the role of metacognitive and cognitive strategies in the acquisition and comprehension of text material. Meryl Baker, in the third presentation, covers the use of computers as tools to investigate individuals' problem solving strategies. In the final presentation, Barbara McDonald discusses the importance of affective variables to cognitive strategies research. Barbara McCombs serves as discussant for the four presentations.

Learning strategies and problem solving training in the military (Ruth Phelps, U.S. Army Research Institute):

In this presentation the use of cognitive strategies in functional environments, such as work settings in the military, is emphasized. After the presentation of a general overview of cognitive strategies, two applied research efforts will be discussed. One effort concerns cognitive strategy training for conducting military intelligence analysis. This training program was developed to improve military personnel's acquisition, retention, and use of information by employing a systematic procedure for organizing their understanding on the basis of a conceptual framework, and subsequently using this framework to solve military intelligence problems. A second effort concerns the use of an interactive videodisc format to teach infantry soldiers problem solving skills. The emphasis in this applicaton is on role-modeling of simple heuristic problem solving strategies across a variety of situations including combat scenarios.

Text processing and metacognitive strategies (Larry Brooks, Doris Bitler, David Shurtleff, U.S. Army Research Institute):

The focus of this presentation is on the use of metacognitive strategies to comprehend text. Metacognitive strategies are those strategies which a learner can use to increase his or her knowledge about how well they understand, or do not understand, new information. For example, in playing a game, a person may not realize that he or she does not know all the rules for that game until the situation arises where that rule is applicable. The research to be discussed in this presentation was conducted to investigate the effectiveness of teaching metacognitive strategies, in conjunction with cognitive strategies such as outlining, as aids in acquiring new information and using that information to solve simple problems. In addition, the role of individual differences (e.g., verbal ability, knowledge of cognitive strategies) in the acquisition and use of metacognitive strategies was assesed. Results and implications of this reseach for military training will be discussed during the presentation.

Generalizable cognitive strategies (Meryl S. Baker, Navy Personnel Research and Development Center):

In this presentation preliminary research concerning the use of computers to collect data on how individuals solve problems will be

discussed. The primary aim of this research is to develop a methodology which allows the researcher to isolate specific processing components involved in the solution of logical problems. This methodology, which is an outgrowth of previous research involving protocol analysis, is one attempt at overcoming some of the problems involved in collecting declarative knowledge from expert problem solvers. The results of this study will be fundamental in developing a tentative cognitive model that incorporates current notions of automaticity into a hierarchically structured set of procedures for solving problems. Results of this study and instructional implications will be discussed during the presentation.

Cognitive and affective processing: Implications for learning strategies (Barbara McDonald, Navy Personnel Research and Development Center):

The emphasis of this presentation is on the relationship between affective and cognitive processing within a training framework. Theories of information processing typically do not specify how cognitive and affective/conative variables interact. The purpose of this presentation is to show that a better understanding of the relationships among cognitive and affective variables will result in more comprehensive theories of cognitive processing, and also provide the educational researcher with a framework within which to formulate strategy training programs. Past research on the relationship between cognitive processing and emotional states has shown that learners tend to recall more information if they are in the same emotional state as when they first acquired the information. Previous work in learning strategies has demonstrated that students can learn to monitor their moods and maintain a positive mood during the acquisition of new knowledge. This typically results in an increase in the amount of information acquired. Preliminary data from the present study will be presented and, based on this data, a tentative model of the interaction between affective and cognitive variables will be proposed.

Discussant (Barbara McCombs, Denver Research Institute):

A critique and discussion of the four presentations and cognitive strategies training in the military will be presented.[1]

[1]The opinions expressed in this paper are those of the authors and do not necessarily reflect official U.S. Army or U.S. Navy policy.

PANEL SESSION

TECHNOLOGY TRANSFER - NATIONAL SECURITY POLICY AND THE SCIENTIFIC AND ACADEMIC COMMUNITY

Cecile Shure Landrum and Patrick McDermott
B-K Dynamics

C. David Hartman
Technology Transfer Intelligence Committee

James Dearlove
Defense Intelligence Agency

ABSTRACT

It wasn't until 1982 during a speech before the American Association for the Advancement of Science and in a Senate Governmental Affairs Subcommittee on Investigations Hearing on Technology Transfer that the Academic and Scientific Community was brought into the public debate concerning the loss of technology to the Soviet bloc. Then Deputy Director of the Central Intelligence Agency, Adm. Bobby Inman made issue that there was a "hemorrhaging" of technology to the Soviets and one of the channels was the U. S. Scientific Community.

Scientists and engineers, on the other hand, are concerned about the control of scientific information by government. Beyond some strictly procedural questions (who reviews and passes judgement on specific publications, for example) there are cultural and value questions at stake or at least perceived to be at stake in dealing with this issue.

1) Academic freedom -- This ranges from concern over constitutional questions regarding the authority of government to impose sanctions on the flow of information to some more practical outcomes in limiting scientists in their communication with one another.

2) Personal identity -- The scientist as a contributor to society has one major product: knowledge. This is the scientist's base of power in a very competitive society. This knowledge contribution is made via publication.

3) Science as an international endeavor -- Control of scientific information for reasons of national security contradities another deeply held sense among many scientists that scientific discovery "belongs to the world." Each scientist's endeavor is thus linked to a historical progression in which generations of scientists build a common knowledge base for the benefit of humanity.

The Scientific Community responded with concern and established a panel on scientific communication and National Security Committee on Science, Engineering, and Public Policy, and headed by Cornell President Emeritus Dale Corson. The purpose of the panel was to "examine the various of the application of controls to scientific communication and to suggest how to balance competing national objectives so as to best serve the general welfare."

The result, in addition to the thoughtful and comprehensive study were a new appreciation and working relationship between the intelligence community and the academic and scientific communities.

Using Comparison-Based Methods for Predicting and Designing

Gary A. Klein Sallie E. Gordon
Klein Associates
740 Wright Street
Yellow Springs, Ohio 45387

Abstract

Effective planning requires that we can predict the outcomes of different options, but it is difficult to make predictions when we have limited knowledge of the relevant causes and their impacts. Comparison-based methods are proposed as a means of generating predictions despite uncertainty and missing data. The methods have the further advantage of providing an audit trail describing the basis for the predictions: the comparison case used, the operational data collected, the adjustments made along with their rationale.

Planning and decision making often require predictions about the consequences of actions. In most cases the outcomes are to be predicted in the context of open systems rather than closed systems. That is, the outcomes are in natural rather than laboratory settings, and they are influenced by a multiplicity of causes, many of which are not well recognized or understood. It is a formidable task to make predictions in the face of considerable ignorance that usually exists about the causal factors affecting the variables we wish to predict.

One hope for a way out of this dilemma rests in the growth of scientific knowledge. As empirical research creates a pyramid of knowledge, then there may become available a firm data base for generating predictions.

Unfortunately, our confidence in empiricism has become weakened. The work of Kuhn (1970) suggested that science may progress through revolutions rather than through the accumulation of linked findings. Each revolution finds a new way to think about problems, and makes previous research appear irrelevant. For fields such as psychology, this would explain why we are not relying very much on data collected from operant conditioning studies with rats and pigeons. Therefore, we may never develop a firm data base in psychology, whereby laws gradually replace theories.

Even more alarming is a recent suggestion by Manicas and Secord (1983) that basic laboratory research may never be able to answer applied psychology questions. Their argument is that scientific laboratory research occurs in a closed system, a laboratory, which is designed to manipulate one causal variable at a time for the purpose of determining whether that variable does indeed have an influence on the dependent variable being measured. However,natural environments contain countless factors that can act together to influence the dependent variable. Scientific procedures are not designed to handle open systems. Scientific knowledge can be used to suggest the causal variables that might be relevant. But that is all that can be expected from scientific research. The rest is up to the applied psychologist who must anticipate how all of the relevant causal factors will interact.

How can we account for the interaction of causal factors when we may not even know what they are or how they work?

It may be instructive to see how this is done in a laboratory context, in order to fashion a strategy for a natural environment.

In the laboratory, the standard method is to randomly assign subjects to experimental and control groups, and make the manipulation of the independent variable the only difference between these two groups. One of the things that this accomplishes is that the control group will reflect all of the other relevant causes affecting the variable that is measured. An experimenter can use the control group as a comparison case, to reflect all kinds of causal factors, whether these are understood or not, or even whether they are completely unknown.

For example, if we think that a certain type of food additive temporarily increases activity levels, we can set up the study and draw conclusions even if we do not know all the factors affecting activity levels, or how any of them work or interact. We can be crushingly ignorant about the causes, and still draw a valid conclusion, because the control group serves as a comparison case to reflect all of the causal factors in the same way that the experimental group does.

In most applied settings, it is not possible to set up a control group that can match on all the relevant causal factors except the one being studied. One thing that can be done is to find a comparison case that can act as a surrogate control group.

This approach was taken by the method of Comparability Analysis, formalized by Don Tetmeyer and his co-workers in the Air Force during 1971-72 as a way of predicting the reliability of new aircraft subsystems. For a given situation, A, if there is a prediction target, A_t, then Comparability Analysis would attempt to identify a comparison case from a similar situation, B. The actual data, B_t, would be collected and used as the prediction for A_t. Now this is clearly too simple-minded since situations A and B are likely to differ in some important ways. Comparability Analysis recognizes this. It uses Subject Matter Experts to estimate the type and magnitude of the differences between target and comparison case, in order to generate an adjustment factor. Thus, the data point B_t is adjusted to reflect the known causal differences and provide a more accurate prediction of A_t.

In one case that we studied, an engineer attempted to predict the reliability of the B-1 hydraulic system in 1976, before any B-1 aircraft were built. He used the B-52 as a comparison system, and obtained the operational data on the B-52 hydraulic reliability. He adjusted these data to reflect pressurization differences between the B-1 and B-52, and obtained his prediction.

Notice what Comparability Analysis is doing. The comparison case is functioning in the place of a control group. The failure to set up a matched control group means that there can be differences in the way that causal factors operate for the target and the comparison cases. An attempt is made to take

this into account by identifying important causal differences and adjusting for them. What about differences for which there are no straightforward adjustments? In some applications, this is enough to reject the comparison case. What about differences involving causes that may not be recognized? It is these "mysterious" causes, outside of our knowledge, that create the biggest problems. In scientific research, these kinds of causes are handled by virtue of the matched control group. For Comparability Analysis, there is also a potential for handling these causes with the comparison case. Thus, we may not know how the flight profile affects subsystem reliability, but using one type of bomber as a comparison for another should generally account for the variable of flight profile.

But what about other causal factors that we do not know about and that operate differently in the comparison and target cases? This is a definite weakness of Comparability Analysis. It is why we would have less confidence in the results of such a study than in a formal laboratory experiment. Comparability Analysis does have a way of addressing this problem because it leaves an audit trail of the rationale behind the prediction. Thus it is possible to determine which comparison case was used, what adjustment was made, and why. The reasoning is explicit, so that someone else may conclude that critical factors were ignored. While this is not a solution to the problem of accounting for unknown causal factors, it is a way of building in a safeguard that may limit the problem.

How valid are the results of a Comparability Analysis? This will depend on the area being studied, the level of effort provided by the Subject Matter Experts, the quality of the operational data that are available. In a study of the predictive validity of the method, Widenhouse and Romans (1977) examined the Comparability Analysis predictions for the A-10 aircraft reliability, as opposed to the actual reliability data after the aircraft completed flight testing and were delivered to the Tactical Air Command. In the derivation of the Comparability Analysis predictions, each aircraft subsystem was compared to a subsystem of an operational aircraft, and the data for that subsystem were adjusted. The major causal factors used in the adjustments were differences in complexity for the target and comparison subsystems; differences in sortie length; and differences in aircraft maturity.*

The original Widenhouse and Romans Technical Report only presented the predicted and actual values, without any attempt to compute correlations. However, we were able to calculate Pearson Product-Moment correlations as a way of measuring predictive validity. For Mean Time Between Failures, the correlation was .76, which accounts for over 57% of the variance. For Mean Maintenance Hours/Flying Hour, the correlation was .84. The validity was lower for predictions of Task Times; two different comparisons yielded correlation coefficients of .46 and .36. Overall, these data show reasonable levels of accuracy. The major problem encountered was that TAC changed the typical maintenance procedure for this aircraft. Previously, new maintenance technicians were assigned when an aircraft was first delivered. For the A-10 the technicians that had worked on it during flight testing were reassigned to stay with it when it was delivered to TAC. This tended to improve reliability over what had been predicted. In defense of Comparability

* The comparison data reflected mature aircraft, which are generally more reliable, whereas the validation data were for initial deployment, when reliability is usually lower.

Analysis, it is not clear that any predictive technique could account for changes in standard procedures unless Subject Matter Experts take note of these changes and their consequences.

Since we were impressed with the potential capabilities of Comparability Analysis as a general predictive method, we attempted to understand the logic behind its application (Klein & Weitzenfeld, 1978; Klein & Weitzenfeld, 1982). Our goal was to be able to improve it and increase its range of application beyond reliability and logistics. We concluded that Comparability Analysis is an application of the inferential process known as reasoning by analogy. The comparison case is the analogue used to generate an inference in a new situation.

Much of the current research on analogical inference is limited in applicability, since it either focusses on one proportion paradigm (A:B::C:D) used in IQ and other tests, or on situations where all the terms of the target and comparison cases map onto each other. Part of the power of Comparability Analysis is that it is designed for situations where the analogy is not perfect. This is fortunate, since the mapping is rarely perfect. Comparabiltiy Analysis includes a means of making adjustments to reflect this imperfection.

We have studied several models of analogical reasoning (Klein & Weitzenfeld, 1978, 1982) including the proportional model examined by Sternberg (1977), the philosophy of science model presented by Hesse (1966), and the similarity matching model of Kahneman and Tversky (1973) before arriving at a model featuring the role of causal factors (Klein, 1982; Weitzenfeld & Klein, 1982). Briefly stated, this model states that for situation A, there are a number of causal factors, (x, y, z....) that together determine the prediction target, A_t. A comparison situation B can be used if it reflects the same causal factors affecting an analogous variable: (x, y, z....) determine B_t. Where we know that there are differences between the levels of a causal variable such as x, we can adjust for these differences. It is the only model of analogical reasoning that we have examined that fully addresses the use of comparison cases for making predictions.

Our analyses have suggested that there are a set of strategies for using comparison cases to generate predictions. We can refer to these collectively as Comparison-Based Prediction. Comparability Analysis is one strategy, a relatively straightforward one. More complex strategies include the use of multiple comparison cases and multiple Subject Matter Experts. One potential problem with a single comparison is that if you have selected the wrong comparison case, you will be mislead. The use of several comparison cases reduces the chances that this will occur. We have also experimented with using a more precise method for identifying the causal factors that differentiate the comparison case and the target case, so that the adjustment factor can be more explicitly tied to specific causal variables.

During the course of our work, we have succeeded in extending Comparison-Based Predictions to domains such as predicting the effectiveness of new training devices (Klein, 1982) and predicting the number of clients who would use a new treatment facility (Klein & Williams, 1983).

We can see some of the advantages of using Comparison-Based methods to make predictions in applied settings. It is designed to handle unknown and

poorly understood causal factors, and missing data. Therefore it is robust, especially compared to formal models. The use of formal models to make predictions can appear to increase predictive validity. However, it is difficult and time consuming to construct these models, and it may not be possible to develop adequate models unless we have a great deal of knowledge about all of the relevant causal variables in a given area. Comparison-based methods appear to be an improvement over simple use of expert judgement, since it generates an explicit audit trail of the rationale behind the prediction. To the extent that it uses experts, it asks them to make relative judgements which seem to be easier to generate than absolute judgements about predicted values.

It is unreasonable to expect that comparison-based methods will always generate perfectly accurate predictions. As the resources and data quality are reduced, validity will be reduced. Comparison-Based Prediction methods have some definite limitations. Nevertheless, it can be argued that for generating predictions under conditions of high uncertainty there are no alternative prediction approaches that can perform as well.

We are currently extending Comparison-Based Prediction methods in the domain of training device effectiveness. We have recently completed a study that attempted to predict the effectiveness of Army maintenance training devices, and we are also in the process of generating predictions about Army training devices for tank gunnery tasks. In addition, we are exploring the use of comparison cases for generating recommendations as well as predictions. In our effort involving maintenance training devices we also were able to use comparisons between different types of trainers to form the basis for recommendations about improved training device design.

The use of comparison cases for generating design recommendations is not entirely novel. Every time a design engineer handles a requirement by searching for a previous application that just needs some modification in order to satisfy the present need, we have an instance of comparison-based design. Usually this is done in an informal manner. Our work may allow us to create an audit trail showing the comparison case used, the areas of mismatch, the modifications proposed and their rationale. This may support the design process and facilitate evaluation and communication about design features.

Based on the research and analysis that we have performed thus far, we are optimistic about the potential for using comparison-based methods to address a variety of Department of Defense needs in the areas of prediction and design recommendations. It is an explicit technique that can be applied to problems where there is incomplete understanding of the causal parameters involved. The methods have limitations. They cannot create knowledge out of ignorance, they cannot extract wise judgement about adjustments from Subject Matter Experts who lack intelligence. They cannot provide guaranteed algorithms for handling unknown causal factors. But these limitations will affect alternative methods as well. Comparison-based methods appear to be an effective way to get the most use out of the existing knowledge for a given problem area.

REFERENCES

(Due to space limitations, references have been omitted from this edition of the paper. References are available upon request from the author at the address given at the beginning of the paper)

The Dual Meaning of the Term Context: Implications for Research, Theory and Applications

Delos D. Wickens
Colorado State University

Abstract

The term context has two psychologically different meanings. The first simply refers to the environment in which some activity takes place; the second to the situation where an event is ambiguous and requires additional information (context) for it to be precisely defined. It is suggested that they be differentiated by adding the suffix alpha to the former and beta to the latter. A literature survey shows context alpha to vary greatly in its effectiveness and reasons for the variation are suggested as well as a theoretical interpretation of how it operates. Context beta is described and related to its underlying psychological mechanisms. Applications of both to Human Factors work are suggested.

During recent years, the term "context" has occurred more and more often in the technical literature of psychology, the concept being used with increasing frequency to account for many behavioral phemonena. If one looks up the word "context" in some standard dictionary, one finds that it has two different meanings. Webster's Third International Dictionary lists the following: (1) the interrelated conditions in which something exists; and (2) the part or parts of a written or spoken passage preceding or following a word or group of words and so intimately associated with them as to throw light upon their meaning.

The first of these definitions, simply refers to the environmental surroundings in which some event exists or occurs, and it contains no implication that the context influences that event. The second definition, states essentially that the context is needed to clarify the meaning of a particular word or group of words, or some action; otherwise the event is ambiguous. Psychological research on context has used both of these meanings somewhat indiscriminately, and, since the underlying mechanisms involved in determining the behavioral outcome differ depending upon which meaning of context is being investigated, it would seem to behoove a science to make a clear conceptual and linguistic distinction between the experimental operations appropriate to each meaning. Hereafter I shall refer to two meanings as context alpha and the second as context beta. I have chosen these suffix terms because they carry no potentially confusing excess meaning.

Context Alpha

Context alpha refers to the situation in which the context--the particular environmental surrounding--is essentially irrelevant to the central task, whose demands and characteristics remains the same regardless of the context. Thus the basic aims of a lecture remain unchanged regardless of the room in which it is delivered; if one must learn a list of French-English vocabulary words, the

pairing of the words in the two languages are constant wherever the list is studied; the pattern of gear shifting in a car is unchanged regardless of the occupants of the car or the local scenery. It is to this type of context that McGeoch in his classic 1932 paper referred when he wrote "the learner is forming associations, not only intrinsic to the material being learned, but also between the parts of this material and the manifold features of the context or environment in which the learning is taking place" (p. 365).

Thus in the instance of context alpha various aspects of the environment present at the time of learning have some likelihood of becoming associated with the learned response and may, at some later time, influence the probability of occurrence of that response. The term "some likelihood" is used because, as research to be reviewed will show, the magnitude of the context effect is quite variable. Logically, context alpha, intrinsically unrelated to the central task as it is, is not a necessary condition for the occurrence of the response. Nor on the other hand should context serve as a sufficient condition for a given behavior. If so, in a conditioning or a monitoring situation where the "true" signals are randomly spaced in time, the intertrial responding or false positives would be so great that even when the response occurred shortly after the onset of the appropriate signal, it could not be unequivocally considered to be a true CR or a true hit.

Experimental designs on context research. The paradigms employed on context research have been addresssed to the two outcomes mentioned above, namely context as a facilitator or inhibitor of performance. One class of experiments falls into the realm of research on transfer of training and the other into the realm of memory, but they are essentially two sides of the same coin. The transfer design requires a minimum of two major groups which are each given equal training on some task in a particular environment (context) and later one group is tested in the same environment and the other in a different environment. I will hereafter refer to these groups as Stay and Change. Typical examples of this type of context alpha manipulation are: Abernathy (1940)--different classrooms; Smith (1979)--different laboratory rooms; Godden & Baddeley (1975)--underwater or open air with scuba divers. In the memory design the two major groups learn two potentially interfering tasks. The Stay group learns both tasks in the same context and recalls one or the other task in that same context. The Change group learns each task in a different environment and the recall test is usually conducted in the context in which the target task was learned. Typical experiments investigating memory and context alpha are Bilodeau and Schlosberg (1951)--different laboratory rooms; Dallett and Wilcox (1968)--marked visual differences; Smith (1982)--different class rooms.

In the transfer design one is concerned with the extent to which the level of performance is degraded by executing the task in an environment different from that in which it was originally learned. In the memory design, wherein two potentially competing activities must be learned, the interest is whether or not the interference can be reduced if the two tasks were learned in a different environment.

The magnitude of the context alpha effect. In this section a few of the many context experiments with both humans and lower animals will be reported in order to demonstrate the wide variation in the effectiveness of the manipulation. This is being done to indicate the need for systematic research identifying the basis of the variation to be reported.

A prototypical transfer study is an experiment by Abernathy. During the regular course in introductory psychology, special lectures were given on such topics as sensory functions or principles of learning, for which carefully standardized tests had been developed. Later the examinations were administered in the same room with the same proctor on in a different but standard classroom with the chairs rotated 180 from normal and with a different proctor. The change in context produced a decrement of only 4 percent, a result that was not significant. Smith (1979) had college students learn a list of words in a pleasantly furnished laboratory and recall them a day later either in the same room or in a sound-proofed chamber. His Stay group performed about 50 percent better than did the Change group. Godden and Baddeley (1975) presented verbal material to scuba divers via auditory communication either while they were underwater or on land and then tested them under a Stay or Change condition. The Stay condition was superior to the Change condition by about 50 percent. The animal literature on context alpha shows an even greater variation in its effectiveness. Bouton and Bolles (1979) established a CER in one chamber and then extinguished it in a very different chamber. When the animals were returned to the original chamber, their CERs did not differ from those of a group that had never undergone extinction. Thomas and McKelvie (1982), using the key peck response of pigeons, found a difference beween the two groups--Stay and Change--of about 45 percent. In summary, the transfer effect has ranged from being infinitely large, as in the Bouton and Bolles work to little, if any, in the Abernathy study, and essentially the same disparity holds in the memory design (Bilodeau & Schlosberg; Wickens, Tuber & Wickens, 1983).

The basis for variation in context alpha effectiveness. It seems obvious that for context alpha to be effective at all the subject must perceive at some minimal level the context change, and furthermore it would seem that the more salient the change the greater the magnitude of the differential in performance beween the Stay and Change groups. Given this latter assumption, one would expect to find a considerable context change resulting from learning underwater and recalling on land or vice versa, but the magnitude of the context alpha effect was about the same in the Godden and Baddeley experiment as in the Smith laboratory research.

The far greater effect of the Change condition in animal compared with human research suggests another reason for the variation. Animal researchers tend to use naive subjects who, after living only in their home cage, are placed in an experimental chamber and given some experimental treatment, usually a shock. It is small wonder that this context becomes so significant to these subjects. Contrast this condition with that of college students who are simply moved to a different classroom. This suggests that the degree of prior experience in many different environments is highly related to context effectiveness, with the more widely experienced individual being less context dependent. From a human factors point of view this would suggest that the degree of similarity of the trainer to the task itself might well differ as a function of the amount of experience of the typical user.

How context alpha operates. In many tasks the subject is required to make different responses when different cue stimui occur, as in a vocabulary list. How can a static environment select the currently correct response; how can context facilitate or interfere with performance? It can do so if its role is to increase the availability of all the responses in the task so that they are more readily triggered when their appropriate cue stimulus occurs. Context alpha can be understood as a mechanism which makes more available the responses which

individually form the class of responses required by that task. In support of this interpretation is the fact that the context effect is absent in recognition memory. This is to be expected since in recognition memory the experimenter supplies the correct responses, depriving context of the role it would otherwise play.

In summary, it is proposed that context alpha has its effect by increasing the availability of the class of response (or single response) which have been acquired in that environment or context. By doing so, the context increases the probabiity that the response will be made when its appropriate cue is presented. This interpretation assumes that the context is not itself discriminating about which one of the multiple responses it has made available will be given at any particular time. That is determined by the cue or stimulus associated with that response. In the transfer design the Stay group is favored by the increased availability produced by the context. But, in the memory situation context makes both classes of response available if they are experienced in the same context. Hence, the Stay group is inferior to the change group.

Context Beta

Context beta refers to the situation in which one stimulus combines with another stimulus event to define the correct response. Each stimulus or stimulus attribute by itself is ambiguous and the concurrence of the two classes of environmental events is a necessary condition for indicating the correct or appropriate response.

Most of the research on acquisition and performance with context beta has been done with intrahumans. These studies indicate that such tasks are more difficult to learn and more likely to produce occasional errors even after they have been fairly well mastered. These results are generally interpreted to be due to the development of both inhibitory and excitatory tendencies toward the specific stimulus since it is correct or incorrect depending upon its current context.

Perhaps the most obvious and important application of the context beta principle to human factors is found in computer usage. The instance is, of course, where a different mode will change the meaning of a particular key or keys.

References

Abernathy, E. M. The effect of changed environmental conditions upon the results of college exams. Journal of Psychology, 1940, 10, 293-301.

Bilodeau, I. M., & Schlosberg, H. Similarity in conditions as a variable in retroactive inhibition. Journal of Experimental Psychology, 1951, 41, 199-204.

Bouton, M. E., & Bolles, R. C. Contextual control of the extinction of conditioned fear. Learning and Motivation, 1979, 10, 445-466.

Dallett, K., & Wilcox, S. G. Contextual stimuli and proactive inhibition. Journal of Experimental Psychology, 1968, 78, 475-480.

Godden, D. R., & Baddeley, A. D. Context-dependent memory in two natural environments: On land and under water. British Journal of Psychology, 1975, 66, 325-331.

McGeoch, J. A. Forgetting and the law of disease. Psychological Review, 1932, 39, 352-370.

Smith, S. M. Remembering in and out of context. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 460-471.

Smith, S. M. Enhancement of recall using multiple environmental contexts during learning. Memory and Cognition, 1982, 10, 405-412.

Thomas, D. R., & McKelvie, A. R. Retrieval of memory in the pigeon by context manipulations. Animal Learning and Behavior, 1982, 10, 1-6.

Wickens, C. D., Tuber, D. S., & Wickens, D. D. Memory for the conditioned responses: The proactive effect of preexposure to potential conditioning stimuli and context change. Journal of Experimental Psychology: General, 1983, 112, 41-57.

AD-P003 366

Research Integration: An Essential for Department of Defense Psychological Research

Laurel W. Oliver
US Army Research Institute for the Behavioral and Social Sciences

Abstract

This paper presents a procedure for integrating the findings of psychological research of interest to the Department of Defense (DoD). In recent years, there has been increasing emphasis on applying more objective, quantitative methods to the integration of research results. This paper reviews various approaches to research integration, describes the meta-analysis approach of Glass (1977), and suggests some directions for the application of meta-analytic procedures to military psychological research.

Introduction

The number of research studies has rapidly increased in many disciplines, including psychology. Within the psychological arena lies a sizable body of research of interest to the Department of Defense (DOD). For decades, a wide variety of psychological research has been conducted in military settings and with military populations (Oliver, in press). Yet this proliferation of research seems not to have advanced the state of science to the extent that one might have expected. In considering the problem of so much research and the relatively few conclusions that can be drawn from it, Frank Schmidt concluded that "the most important problem in psychology and the social sciences today is the failure to produce cumulative knowledge" (Schmidt, 1980).

Accordingly, the purpose of this paper is to review various approaches to research integration, to describe in more detail the meta-analysis approach of Glass (1977), and to suggest some directions for the application of meta-analytic procedures to military psychological research.

Approaches to Research Integration

In this paper, research integration refers to combining the research results of a group of studies. Several procedures that can be used to integrate research findings are described below.

---

[1] The views expressed in this paper are those of the author and do not necessarily reflect the views of the US Army Research Institute or the Department of the Army.

Literary approach. The literary or narrative approach has been the traditional procedure for integrating research findings. The reviewer reads the studies on a given topic and attempts to derive generalizations about the outcomes of that body of research. This can be an extraordinarily difficult task, especially if the number of studies is large or the results conflict. In addition, due to the subjective nature of this process, different reviewers may reach different conclusions when integrating the results of essentially the same body of research (Smith, Glass, & Miller, 1981). Such reviews typically end with a call for more research to resolve the conflicts.

Box-score approach. In the box-score or vote-counting approach, a directional hypothesis is assumed. The reviewer classifies the findings of each study into positive significant, negative significant, and nonsignificant categories and tallies the results. (Reviewers usually do not indicate the direction of the nonsignificant results.) Again, the results may conflict and make it difficult for the reviewer to draw firm conclusions about the research being integrated. One objection to the box-score approach is that it does not take into account the magnitude of the significant results. Hedges and Olkin (1980) have also demonstrated that the probability is high of failing to conclude there is a positive effect when in fact there is and also that the probability of making this error may increase as the number of studies increases.

Meta-analysis. According to Glass (1977), meta-analysis is the statistical analysis of the analytic results from a number of independent studies. To accomplish such an analysis, the findings of the individual studies must be quantified. The approach developed by Glass and his colleagues (Glass, McGaw, & Smith, 1981) is described in the next section of this paper. However, there are other, related meta-analysis procedures. Hunter, Schmidt, and Jackson (1982) have expanded their validity generalization techniques into a full-fledged meta-analytic approach. Hedges and Olkin (1980, in press) have advanced statistical theory and applications to meta-analysis. Also, Rosenthal (1978) has summarized a number of methods for combining the probabilities obtained from the result of two or more studies.

The Meta-Analysis Approach

The meta-analysis procedure most frequently employed to date (and the one which will be described below) has been the one developed by Glass and his colleagues. The unit of analysis in this approach is a standardized mean difference called the effect size.

To calculate the effect size. In their meta-analysis of psychotherapy research, Smith and Glass (1977) (see also Smith, Glass, & Miller, 1980) have defined the effect size as the difference between the means of the experimental and the control groups on a given dependent variable divided by the standard

deviation of the control group. That is,

$$\text{Effect Size} = \frac{M_E - M_C}{SD_C}$$

Using this formula, it is possible to express the standing of the average experimental group subject in terms of the control group distribution (see Smith & Glass, 1977; Spokane & Oliver, 1983). As McGaw and Glass (1980) warn, effect size is a simple concept, but its calculation can be complicated by differences in experimental design and the metric used. These authors suggest using the control group standard deviation as the denominator (McGaw & Glass, 1980, pp. 106-123). Hunter et al. (1982) agree with Glass that the experimental group treatment may affect the experimental group standard deviation as well as the experimental group mean. However, they argue for using the within-group standard deviation because the control group standard deviation has more sampling error and also because research reports are likely to contain values for *t* and *F* and fail to report the control group standard deviation (Hunter et al., 1982, p. 101).

Reports lacking means and standard deviations. If the data (means and standard deviations) are not reported in the study in question, it may be possible to estimate or retrieve the required data from those statistics that are reported (such as *t* and *F*). Procedures are also available for calculating effect sizes from correlation coefficients, nonparametric statistics, and dichotomous outcome variables (Glass et al., 1981). Occasionally, data can be obtained from the author of the article or report, but this procedure is time-consuming and generally unproductive (Oliver & Spokane, 1983).

Sometimes it is impossible to retrieve enough data (either from the research report or the author) to calculate effect sizes (Oliver & Spokane, 1982; 1983). The inadequacies of research reporting were dramatically apparent recently when the author and a colleague surveyed the research on senior leadership to determine the feasibility of using a meta-analysis approach on this literature. Of the 64 studies previously identified as empirical studies of senior leadership (Kimmel, 1981), only about 10% were suitable for use in a meta-analysis. This was a disappointing outcome, for senior leadership is currently of great interest to the Army. It had been hoped that a quantitative synthesis of the empirical findings in the area would be useful to the leader development research program being conducted by the Army Research Institute.

Suggestions

Following are some suggestions for capitalizing on research already accomplished and providing direction for future research in the DoD.

The meta-analysis approach has been used to integrate research findings in dozens of topical areas, primarily in the disciplines of psychology and education. It is important that such efforts be continued in order to determine what we can confidently conclude from our research as well as to identify gaps in our knowledge. In fact, it is likely that journal editors eventually will require that a meta-analysis approach be used for all literature reviews.

Until now, most research integration has concerned the civilian sector. These findings will have their greatest applicability for DoD civilian research. Little quantitative integration seems to have been accomplished for research relating to military populations and settings. At present, we are not certain to what extent findings based on civilian populations apply to the military. It will be important to use the meta-analysis approach to integrate the military research and then to compare the resulting findings with those of research conducted in the civilian sector.

References

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. Review of Research in Education, 5, 351-379.

Glass. G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Hedges, L. V., & Olkin, I. (1980). Vote-counting methods in research synthesis. Psychological Bulletin, 88, 359-369.

Hedges, L. V., & Olkin, I. (in press). Statistical methods in meta-analysis. New York:

Hunter, J.E., Schmidt, F. L., & Jackson, G. B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Kimmel, M. J. (1981). Senior leadership: An annotated bibliography of the military and nonmilitary literature. (Technical Report 532). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

McGaw, B., & Glass, G. V. (1980). Choice of the metric for effect size in meta-analysis. American Educational Research Journal, 17, 325-337.

Oliver, L. W. (in press). Military psychology. In R. Corsini (Ed.), Encyclopedia of psychology. New York: Wiley.

Oliver, L. W., & Spokane, A. R. (1982). Sufficiency in the reporting of research results: Some guidelines. (Technical Report 559). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Oliver, L. W., & Spokane, A. R. (1983) Research integration: Approaches, problems, and recommendations for research reporting. Journal of Counseling Psychology, 30, 252-257.

Rosenthal, R. (1978). Combining results of independent studies. Psychological Bulletin, 85, 185-193.

Schmidt, F. L. (1980, September). The research task of the 1980s: Integrating research findings across studies to produce cumulative knowledge. Paper presented at the meeting of the American Psychological Association, Montreal.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Baltimore, MD: John Hopkins University Press.

Spokane, A. R., & Oliver, L. W. (1983) The outcomes of vocational intervention (pp. 99-138). In W. B. Walsh & S. H. Osipow (Eds.), Handbook of vocational psychology. Vol. 2: Applications. Hillsdale, NJ: Lawrence Erlbaum Associates.

What Training Is Depends on Who You Ask

D. M. Kristiansen, B. G. Witmer,
B. A. Black, and S. L. Burroughs
US Army Research Institute, Fort Knox Field Unit

Abstract

Repeated in-depth interviews were conducted with all levels of the chain of command in two tank battalions in Europe during a training effectiveness analysis of a tank system. The purpose of the interviews was to describe the state of training in the battalions across a six month time frame. Interviews were conducted with the battalion commanders, all the company commanders, over 80% of the platoon level leaders (platoon leaders, platoon sergeants, and tank commanders), over 70% of the crewmen, and over half the mechanics, and maintenance supervisors. The resulting descriptions of the amount of training conducted/received and the nature of that training differed remarkably at three levels of the chain of command when training was rigorously defined and when training was undefined. Implications for data collection and interpretation of field data are drawn.

In October 1982, the Army Research Institute Field Unit at Fort Knox (ARI) conducted in-depth interviews with all levels of the chain of command in an armor battalion in Europe. The research objective was to pilot a method and prototype data collection format for describing the state of training in an operational unit. Project personnel were in the battalion for two weeks which permitted the collection of a large sample: the battalion commander, all the company commanders and first sergeants, over 80% of the platoon level leaders (platoon leaders, platoon sergeants, and tank commanders), over 70% of the crewmen, the battalion maintenance supervisors, and over half the mechanics. Out of these interviews ARI personnel were able to document, in considerable detail, the state of training in the unit as seen by unit members at various points in the chain of command.

How to define "training" for these interviews became an issue during the planning phase. It was decided that a relatively restrictive definition was preferable because it would exclude those activities on the unit training schedule that were mission operations, that is, details, inspections, the conduct of maintenance, physical training, etc. Accordingly, training was defined as "an event whose purpose was to train MOS or Soldier's Manual tasks."

When the team arrived in Europe, the logistics of collecting the data led to a change in the original planning. The ARI researchers did not conduct the interviews of company commanders and first sergeants. These interviews were conducted by members of an accompanying organization. The ARI researchers, remembering long discussions with members of the accompanying organization

concerning the definition of training, assumed that these interviewers used the same definition.

All respondents were asked how much training had taken place in the two week period preceding the data collection period. Company level leaders (commanders and first sergeants) were asked, on the interview forms, to describe training in three different ways: they were asked to estimate the percentage of scheduled training that was actually conducted, to estimate the number of hours of scheduled, unscheduled, individual, and collective training that had actually been administered, and the number of hours of training that they had personally monitored or supervised.

Table 1 displays the results of this line of questioning for scheduled training. The estimates of training conducted were substantially higher among company level leaders than among platoon level leaders and crewmen. Estimates of unscheduled, individual, and collective training showed similar patterns.

TABLE 1

AVERAGE ESTIMATED NUMBER OF HOURS OF SCHEDULED TRAINING

ADMINISTERED	COMMANDERS	43
	FIRST SERGEANTS	32
CONDUCTED	PLATOON LEVEL LEADERS	0
RECEIVED	CREWMEN	1

Table 2 shows the percent of respondents that said that "no" scheduled training (zero hours) had been conducted by platoon level leaders or received by crewmen. This discrepancy between commander's estimates and the estimates of platoon level leaders and crewmen could have arisen in two ways; commanders really did have higher estimates or the commanders were including activities in their concept of training that crewmen were not.

TABLE 2

PERCENT OF RESPONDENTS INDICATING THAT "ZERO" HOURS OF TRAINING HAD BEEN:

CONDUCTED	COMMANDERS	0
	PLATOON LEVEL LEADERS	75
RECEIVED	CREWMEN	80

When the added interviewers were queried after data collection, the ARI team discovered that they were not aware of the restrictive definition of training and, hence, did not define "training" for commanders and first sergeants. This suggested that the differences between the estimates of training were differences in kind rather than differences in magnitude.

The discrepancy appeared to be a measure of the difference between two conceptions of what constitutes training. When training was not defined, company level leaders were quite willing to include any and all activities on the training schedule (for example, guard, drill and ceremonies, maintenance operations, and other mission operations) as training because they felt that all these activities had training value. Training, for the company level leaders, was something that the unit did; it was a "unit" activity.

At the bottom of the chain of command, however, training is what happens <u>to</u> "individuals." The platoon level leaders and crewmen reported that, essentially, no training in MOS or Soldier's Manual tasks had taken place. Many felt they were not being trained in skills they needed and said that they wanted this training.

Realizing that training in garrison is difficult, and aware of the commonly held opinion that "real" training takes place in the field during exercises, questions were asked regarding training during a recent major field exercise. Unit leaders talked, in general terms, of how much more training was conducted during such exercises as compared to what could be accomplished in garrison. Crewmen had mixed opinions, however. The training that a crewman received depended on what duty position he held. Tank commanders and drivers were occupied, from time to time, when the exercise required them to displace from one position to another. Loaders and gunners, however, went along for the ride. They had nothing to do. No "in the cracks" training took place. Leaders were, of course, totally occupied during the exercise.

Leaders appeared to assume that when they were busy, their men were also busy. By crewmen reports, this was just not the case. In these battalions, leaders had little opportunity to see for themselves what the implementation of the training schedule looked like "on the ground." They were pinned down by an abundance of actions requiring the attention of a commissioned officer, often the commander himself.

The other side of the training coin is, of course, mission activities. Most everyone agreed on what was taking place in the unit during the visit; maintenance, inspections, and details. Unit leaders again differed from crewmen on the extent to which they saw these activities consuming duty time. Company level leaders saw maintenance and details as all consuming missions. The team was continually directed to the motor pool because "everyone is down there working on the tanks." Often, when the team arrived in the motor pool, it was essentially empty. For the unit leaders, maintenance was a vital mission. The unit was standing down from a major exercise and preparing for a major inspection. The leaders were concerned with insuring that the unit passed the inspection. They spent much of the two week period "managing the process of" preparing for this inspection. Requirements from higher headquarters and routine actions requiring their personal attention kept them from doing much "in the motor pool" supervision. Table 3 shows the extent to which respondents reported being involved in maintenance, inspections, and details. Entries in the Table are the average number of responses within respondent and mission categories. Platoon level leaders who were closer to what was actually happening in the unit reported somewhat less involvement in maintenance and details. Crewmen, who were the performers of these maintenance and detail activities, reported still less involvement.

TABLE 3

NON-TRAINING ACTIVITIES
(AVERAGE NUMBER OF RESPONSES)

	CREWMEN	PLT LEVEL LEADERS	CO LEVEL LEADERS
MAINTENANCE	0.7	1.2	2.0
INSPECTIONS	0.7	0.8	0.2
DETAILS	0.3	0.5	0.8

This data collection experience suggests that "training conducted" and "training received" are quite different. If one has the task of determining, by questionnaire or interview, what training activities are occurring in a unit, a careful definition is in order. Had the team not defined training carefully for platoon level leaders and crewmen, they may have come away with an interpretation vastly different from what they did. They certainly would not have found out that no training in MOS of Soldier's Manual tasks was taking place.

Second and third visits to Europe in 1983 provided an opportunity to determine if these differences were really differences in kind or magnitude. This actually became necessary because the definition of training used by ARI in the October 1982 visit came under considerable criticism by military personnel. They objected to the restrictive definition of training and asked that, during a visit to the battalion and a sister unit in April 1983, ARI purposely did not define training for respondents. The idea was to see if the results obtained during the first visit were dependent on training being rigorously defined. There was a feeling that if the term "training" were left undefined, the responses given across different levels of command would be more homogeneous.

Table 4 shows that, essentially, the same results were obtained. Commanders reported more time devoted to training than platoon level leaders who, in turn, reported more training than crewmen. The differences are not as marked but still suggest different views of the training world at two relatively close points in the chain of command.

TABLE 4

HOURS DEVOTED TO TRAINING
(LAST TWO WEEKS)

	CDRS	PLT LEVEL LEADERS	CREWMEN
UNIT SPENT IN TRAINING	25		
SPENT SUPERVISING TRAINING	20		
SPENT INSTRUCTING		16	
SPENT IN TRAINING			13

Representatives of the different levels of the chain of command also differed in what training they said had taken place during preparation for a major tank unit exercise - Level I Gunnery. For example, Table 5 shows that commanders reported that their units had conducted Tank Crew Proficiency Course and Tank Crew Gunnery Skills Test exercises prior to Level I Gunnery. Few crewmen reported taking part in such exercises. Observations during the conduct of a Tank Crew Proficiency Course illustrated the problem. The company commander and all six platoon leaders and platoon sergeants were present at the exercise. Fewer than half the crewmen in the unit were present, however, because of guard duties and other details and the normal, day-to-day functions that pull men from a unit.

TABLE 5

PERCENT OF RESPONDENTS REPORTING TRAINING
EXERCISES IN PREPARATION FOR LEVEL I GUNNERY

	CDRS	PLT LEVEL LEADERS	CREWMEN
NO TRAINING	0	17	22
TANK CREW PROFICIENCY COURSE	87	29	13
TANK CREW GUNNERY SKILLS TEST	62	23	0
MINI-TANK RANGE	37	8	0
CREW DRILLS	37	8	0

The implications of these findings seem clear. The perspective of the respondent must be clearly defined when designing interview formats and interpreting interview data and the perspective of one level of the chain of command cannot be assumed for any other level. Data collection at only one point in the chain of command, or at two points closely related (i.e., battalion and company), will yield a biased picture of what is actually going on. Commanders respond in terms of unit involvement. Unit involvement does not necessarily imply the involvement of sizable numbers of men. Soldiers respond in terms of what has happened to them. Both perspectives are necessary to

accurately describe the state of training in a unit. And, of course, the interview referent must also be carefully defined (or undefined) for all levels of the chain of command.

AD-P003 368

Video Games:

A Human Factors Guide

to Visual Display Design and Instructional System Design

Douglas J. Bobko

US Army Research Institute

Abstract

Electronic video games have many of the same technological and psychological characteristics that are found in military computer-based systems. For this reason, the video game is both a fascinating object of study and a valuable experimental apparatus. The results of two on-going research programs, both of which employ video games as experimental stimuli, are presented here.

The first research program seeks to identify and exploit the characteristics of video games in the design of game-based training devices. The second program is designed to explore the effects of electronic video display characteristics on perceptual judgments. The empirical results of these two programs are shown to have practical application in training device design and visual display design.

Video Games and Training Device Design

Although games in general have been employed as learning vehicles, the especially compelling characteristics of electronic video games have not been fully explored for possible exploitation in the design of computer-based training systems. In fact there is little empirical evidence regarding the perceived or salient qualities of these new electronic video games. Accordingly, the following experiment was designed to determine those dimensions along which electronic video games are perceived to vary.

Method

Twenty right-handed male first-year college students served as subjects. The experimental stimuli consisted of ten commercial electronic video games and an "ideal" video game. The "ideal" video game was defined to subjects as their notion of the best or ideal video game, whether imaginary or real.

The twenty subjects were asked to make pairwise comparisons of the similarity of the eleven video game stimuli. Judgments were made with the use of rating scale in which 0 was defined as "these two games are not at all similar" and 100 was defined as "these two games are exactly alike." All judgments were made in a commercial video game arcade closed to the public.

Results and Discussion

The game similarity data were subjected to a multidimensional scaling (MDS) analysis. Given judgments or measures of similarity, MDS provides a spatial representation of stimuli that reveals dimensions relevant to an observer. For example, if two stimuli are judged to have a similar quality, they are placed near each other in a multidimensional space; stimuli judged to be dissimilar are located distant from each other.

A three-dimensional MDS solution was deemed appropriate after examining the squared correlations (RSQ) associated with solutions for different dimensional spaces. The RSQ is a measure of the proportion of variance of monotonically transformed data accounted for by the MDS model. The RSQ's for a one-dimensional (1D) solution through a five-dimensional (5D) solution for the game similarity data are: 1D-0.225; 2D-0.294; 3D-0.314; 4D-0.289; 5D-0.292. Here, the RSQ decreases at the 4D solution and clearly indicates that no more than three dimensions should be considered for interpretation. This decrease in RSQ reflects the subjects' insensitivity to differentiating the games on more than three dimensions.

The first dimension, called "Destructiveness", orders games according to the type of response stragegy necessary for successful play. One end of this dimension is anchored by games which involve the destruction of alien objects by the proactive manipulation of a "mother-ship" capable of firing bullets/lasers/etc. The other end of this dimension is anchored by games which require reactive, avoidance strategies (e.g., avoidance of ghosts, alligators, or barrels). Thus, the first dimension ranges from proactive destruction to reactive avoidance.

The "ideal" game was positioned near the center of the "Destructiveness" dimension. An examination of the raw similarity scores indicated that subjects fell into two groups with respect to the characteristics of an ideal game on this dimension. Specifically, destruction of opponents is an ideal characteristic of video games for half the subjects; the other half prefer games which require avoidance strategies.

The second dimension, called "Dimensionality", locates games according to the number of physical dimensions (on the video screen) in which the player can maneuver. On one end of this dimension, the player's "piece" can be effectively moved in only one dimension (left to right on the visual screen). In the games near the center of this dimension, player movement is two-dimensional (left-right and up-down). Finally, the games on the opposite end provide the visual appearance of movement in three-dimensional space.

It is noteworthy that most subjects found simulated three-dimensional games quite difficult to master and that, at the same time, the "ideal" game is placed between the two- and three-dimensional game groupings. This phenomenon is consistent with so-called "optimal-level" theories of motivation which postulate that moderate levels of cognitive/physical complexity are maximally arousing. Hence, games which allow movement in only one dimension may be too "simple" and three-dimensional games too "complex", in so far as these games are presently configured. Two-dimensional games (or something "between" two- and three-dimensional games) may be maximally motivating. As experience or familiarity with these games increases, one would expect game

preference to shift toward the more "complex" or "difficult" ones.

The third dimension revealed by MDS, "Graphic Quality," orders games by their degree of color vibrancy and resolution. At one end of the continuum, games are characterized by a relative lack of color and by relatively low graphic resolution. At the other end, games are characterized by vibrant colors, fine detail, and high graphic resolution. On this continuum the "ideal" game is located at the graphic quality extreme, suggesting that subjects prefer games with high quality graphics.

Applications and Future Research

These results have several applications in the design of computer-based video training systems. For example, the findings suggest that video game-based lessons will be more compelling if they are designed with high quality graphic displays and simulated movement in two or three dimensions. A second application derives from the finding that some subjects prefer avoidance to destructive behaviors (i.e., reactive to proactive behaviors). Implied here is the need for a fit in personnel training between this cognitive preference and the type of task required for successful completion of a game-based training lesson. Consider, for example, a mathematics lesson embedded in a game where correct answers result in the avoidance of falling barrels. Such a lesson might be most compelling (ideal) for those subjects who prefer games which require reactive, avoidance behaviors. Conversely, subjects who prefer proactive game behaviors would find the same mathematics lesson most compelling if correct answers result in the destruction of alien objects.

Finally, the present results may be recast in an individual differences framework. That is to say, the psychological predispositin toward destructive (or avoidance) game behaviors may be a valid instrument for classifying personnel and predicting future success. This classification would be of special use when, say, the task to be learned is inherently destructive (e.g., shooting enemy aircraft) and difficult to embed in a game-based training device that required avoidance behaviors. In such a situation, subjects who prefer avoidance game behaviors would not be selected for training. Hence, future research should consider the role of personality variables as a moderator of game-based training success.

Video Games and Visual Display Design

The modern military environment is increasingly characterized by the electronic exchange and display of large amounts of information. It is important to determine whether and understand how the characteristics of these visual information displays affect perceptual judgments and consequent decisions.

The effect of visual display characteristics on the perception of time has received little attention. However, the accurate perception of time is of increased importance in job environments where decisions are time-critical and where information is received via electronic visual displays. These implications are best highlighted by means of a hypothetical example:

The commander of an armor battalion sits inside a mobile command post several miles behind the forward edge of the battle area. The battalion commander is receiving information about the battle on a computer screen which simulates the terrain and troop strength of his forces and the enemy forces; the screen is 4 inches on its diagonal. He finds that his position is under heavy conventional weapons attack and that the enemy is close to overrunning his defenses. Division command informs him that reinforcements will not arrive for another 30 minutes. He is further instructed to hold his position no matter what the cost and he is authorized to use a theater nuclear weapon if the defensive line is broken.

Will the commander misjudge the passage of time because of the characteristics of the visual display? One characteristic of interest is the scale of the information presented on the visual display. Accordingly, the purpose of the following experiment was to determine whether the perception of time is affected by the scale of the visual display.

Method

The subjects were 72 first-year college students, half of whom were female. The independent variable for this experiment was the screen size of a television monitor which was connected to a commercially available video game. There were three levels of screen size: a 5 inch (on the diagonal) screen, a 12 inch screen, and a 23 inch screen. The subjects were randomly assigned to one of the three screen size levels with the restriction that half the subjects assigned to each screen size were female.

The subjects were asked to play a computer video game which displayed a simulated combat situation on the television screen. The object of the game was to shoot and destroy moving enemy targets by using a joystick to aim and fire a gun.

Each subject was given 55 seconds of playing time. At the end of this interval, the subject was asked to estimate (in seconds) as accurately as possible the length of time the simulated combat situation had lasted. These time estimates served as the dependent variable for the experiment.

Results and Discussion

A 3 x 2 analysis of variance was performed on the time estimates. There were three levels of screen size: 5 inch, 12 inch, and 23 inch. There were two levels of the second variable, sex of the subject.

The effect of screen size on time estimates was found to be significant ($p = .005$). The mean time estimate for the 5-inch screen was 120.5 seconds; the 12-inch screen mean time estimate was 76.1 seconds; and the mean estimate for the 23-inch screen was 67.9 seconds. In other words, the perception of elapsed time for a given clock interval (55 seconds) increased as screen size decreased.

The effect of sex on time estimates was significant at $p = .06$, and females tended to give longer time estimates at each level of screen size. The interaction of screen size and sex did not approach significance.

Applications and Future Research

The preliminary results of this research clearly suggest that the scale or size of a visual display affects the perception of time. The most pronounced "distortion" of time perception was found with a 5-inch visual display -- this is typical of the size of visual monitors found in aircraft, ships, and tanks.

Because this research is in its early stage, the scope and generality of the findings for the design of visual displays is not certain. Nevertheless, the results reported here have potentially dramatic significance for environments in which information is visually displayed and in which decisions are time-critical.

Summary

The popularity of electronic video games is of interest because these games have qualities similar to those found in various military computer-based systems. Moreover, the commercial video game arcade is an *in vivo* proving ground for the "fittest" of electronic games -- hence these games may provide design guidelines for system developers.

The exploratory and preliminary results discussed above clearly demonstrate that the video game is both a valuable object of study and a useful experimental apparatus. The results also serve to define the direction of video game research now underway.

Footnote

The views, opinions, and/or findings contained in this report are those of the author and should not be construed as an official Department of the Army or Department of Defense position, policy, or decision, unless so designated by other official documentation.

Requests for further information regarding this research should be sent to Dr. Douglas J. Bobko, US Army Research Institute, ATTN: PERI-II, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

AUTHOR INDEX

U.S. GOVERNMENT PRINTING OFFICE 1984--778-261/9190 REGION NO. 8